### 6475789

+

1/34

MOCLIF O AKFLHWLMSVYVVELLRSFFYVTETTFQKNR ISEIEWLVLGKRSNAKMCLSDFEKRKQIFAEFIYWLYNSFIIPILQSFFYITESSDLRNR LKDFRWLFISDIWFTKHNFENLNQLAICFISWLFRQLIPKIIQTFFYCTEISSTVT- TREISWMQVET-SAKHFYYFDHEN-IYVLWKLLRWIFEDLVVSLIRCFFYVTEQQKSYSK	MOCIÍ 1 LFFYRKSVWSKLQSIGIRQHLKRVQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGL TVYFRKDIWKLLCRPFI-TSMKMEAFEKINENNVRMDTQK-TTLPPAVIRLLPKKNTF IVYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLSNFNHSKMRIIPKKSNNEF TYYYRKNIWDVIMKMSI-ADLKKETLAEVQEKEVEEWKKS-LGFAPGKLRLIPKKTTF . * * * * . * . * . * . * . * . * . * .	MOLİÉ 2 RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLNYERA RLITN-LRKRFLIKMGSNKKMLVSTNQTLRPVASILKHLINEESSGIPFNLEVYMKLLTF RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKEF RPIMTFNKKIVNSDRKTTKLTTNTKLLNSHLMLKTLKN-RMFKDPFGFAVFNYDDVMKKY	MOCIÍ 3 (A) KKDLLKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPEFVIRKYATIHATS KQRLLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN EEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKN
human tez1 EST2 p123	human tez1 EST2 p123	human tezl EST2 p123	tez1 EST2 p123

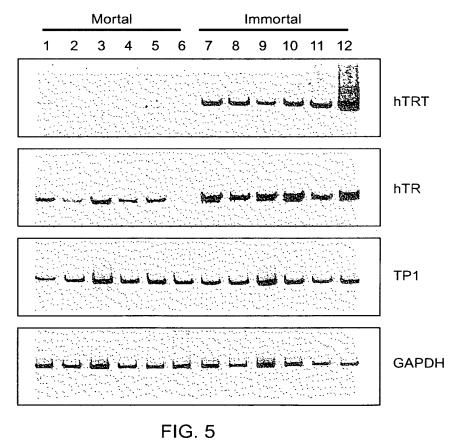
FIG. 1

•	-	1 2 A	я. П		Mol. weight	ā
<b> </b> ∢				<ul><li>Sp_Trt1p</li></ul>	116.000	10.6
				——hTRT	127.000	11.3
				—— Ea_p123	123.000	10.1
ו מע נ				Sc_Est2p	103.000	10.0
	msDNAs					
	Mito.plasmid/RTL					
	Group II introns					
	Non-LTR Retrotransposons					
	Hepadnaviruses					
	LTR Retrotransposons (Copia-Ty1)		<b>모</b>			
	LTR Retrotransposons (Gypsy-Ty3) 路路化ZZI	NATIONAL NA				
	Caulimoviruses	MATCH NAME OF THE PROPERTY OF				
	Retroviruses	MANAGE OF THE PROPERTY OF THE				
		HIV-1 RT				

FIG. 7.

+

PALM THUMB **FINGERS** FIG. 3



4/34	181 197 146
82 87 87 68 20 25	
hhK K TERIVKKKLKD LTEVIASIIKP LSTFLKTTKLL MRILKDALKN IKELKRYISD LIKELKRYISD	Motif E  W G S  17 KKRMPFFGFSV  19 HGLFPWCGLLL 23 QDYCDWIGISI 20 KELEVWKHSST  hLG h  4 ETPARFLGVIL 25 ESKGYLGVIL 0 EPPFLWMGITL
Motif A  PCLYFh hDh CYD I hhK K 40 FGRKKYFVRIDIKSCYDRIKQDLMFRIVKKKLKD 45 PPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKP 41 GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL 42 VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN AF  h hDh GY h 27 FGGSNWFIEVDLKKCFDTISHDLIIKELKRYISD 32 RKEYCSAVFLDISEAFDRVWHEGLLLKLAKILPY 0 LKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIP	Motif C  LLRL DDFLhIT  6 LLRVVDDFLFITVNKKD 0 AKKFLNLSLRGFEKHNFSTSLEKTVI 1  5 LLRLVDDFLLVTPHLTH 0 AKTFLRTLVRGVPEYGCVVNLRKTVV 1  14 LMRLTDDYLLITTQENN 0 AVLFTEKLINVSRENGFKFNMKKLQT 2  8 ILKLADDFLISTDQQQ 0 VINIKKLAMGGFQKYNAKANRDKILA F  F  h Y DDhh  55 YVRYADDILIGVLGSKN 2 KMIKRDLNNFLNS-LGLTMNEEKTLI 7  7 LSTYADDTIVLSSDILA 6 NENYLKTFSDWADKWGISVNAAKTGH 2  4 IYQYMDDLYVGSDLEIG 1 HRTKIEELRQHLLRWGLTPDKKHQK
## Motif 2  FRh h h K K  O FRLYFh hDh CYD I hhK K  O FRLITNLRKRFLIKMGSNKKMLVSTNQTL 40 FGRKKYFVRIDIKSCYDRIKQDLMFRIVKKKLKD  O LRPIVNMDYVVGARTFRREKRAERLTSRV 45 PPPELYFVKVDVTGAYDTIPQDRLTEVIASIIKP  O FRPIMTFNKKIVNSDRKTTKLTTNTKLIN 41 GQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLL  I FRIIAIPCRGADEEFTIYKENHKNAIQP 42 VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN  R h K  N h Dh GY h  O IRPLSVGNPRDKIVQEVMRMILDTIFDKK 27 FGGSNWFIEVDLKKCFDTISHDLIIKELKRYISD  O IRPLSVGNPRDKIVQEVMRMILDTIFDKK 37 FGGSNWFIEVDLKKCFDTISHDLIIKELKRYISD  O IRPLSVGNPREKLLKRRFRVDLFK 32 RKEYCSAVFLDISEAFDRVWHEGLLLKLARILPY  WRKLVDFRELNKRTQDFWEVQLGIPHPAG 0 LKKKKSVTVL_DVGDAYFSVPLDEDFRKYTAFTIP	Motif C LLRL DDFLhIT 6 LLRVVDPELFITVNKKD 5 LLRLVDPELLYTPHLTH 14 LMRLTDDYLLITTQENN 8 ILKLADPELISTDQQQ F h Y DDhh 55 YVRYADPILIGVLGSKN 7 LSTYADPILIGVLGSKN 7 LSTYADPILIGVLGSKN 7 LSTYADPILIGVLGSKN 7 LSTYADPILIGVLGSKN 7 LSTYADDILYVGSDLEIG
h hRhiPKK p hNVRMDTQKTTLPPAVIRLLPKKNT- EVRQHREARPALLTSRLRFIPKPDG- KEVEEWKKSLGFAPGKLRLIPKRTT- CRNHNSYTLSNFNHSKMRIIPKKSNN p hh h K LSNELGTGKFKFKPMRIVNIPKPKGG SILRIGYYPDAWKHAQVKMILKPGKS EGKISKIGPENPYNTPVFAIKKKDST	Motif B'  TRT con K Y Q GIPQGS LS hL h Y DL  SP Trtlp SQYLQKVGIPQGSILSFLCHFYMEDLIDEYLSFT  HTRT KSYVQCQGIPQGSILSFLCSECYGDMENKLFAGI  Ea_pl23 KFYKQTKGIPQGLCVSSILSSFYYATLEESSLGFL  Sc_Est2p KCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFK  RT con hPQG pP hh h  Sc_a1 TYHKPMLGLPQGSLISPILCNIVMTLVDNWLEDYI  Dm_TART RAGQIGAGVPQGSNLGPILYSIFSSDMPLPHIYHP  HIV-1 GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFKKQN
TRT con Sp rrtlp hTRT Ea p123 Sc Est2p RT con Sc a1 DM TART HIV-1	TRT con Sp Trtlp hTRT Ea p123 Sc Est2p RT con Sc a1 Dm TART

WL hh hh pery te p y rk w L h i K 729 WLYNSFIIPILQSFFYITESSDIRNRTVYFRKDIWKLLCRPFITSMKM 8 546 WLMSVYVVELLRSFFYVTETFQKNRLFFYRKSVWSKLQSIGIRQHLK 10 441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSIADLKK 8 366 WLFRQLIPKIIQTFFYCTEISSTVT-IVYFRHDTWNKLITPFIVEYFK 8

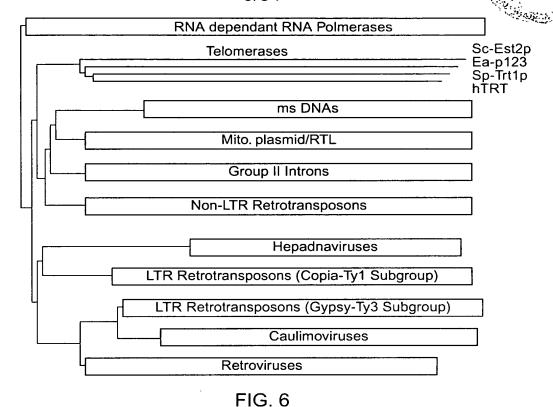
TRT con Sp\_rrtlp hTRT

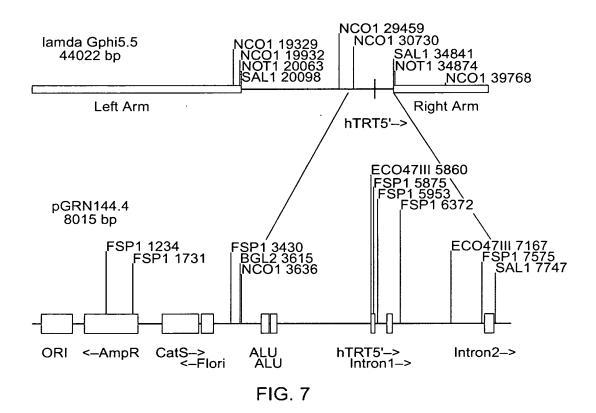
Ea\_p123 Sc\_Est2p

Motif T

+

FIG. 4







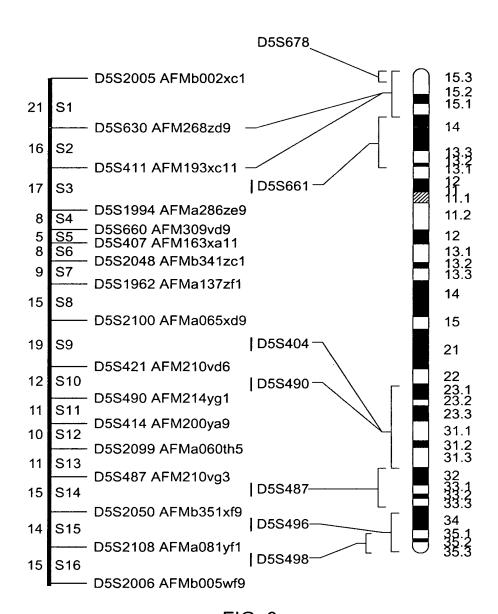
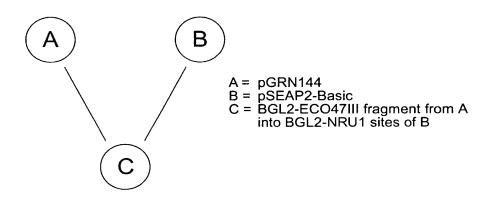


FIG. 8





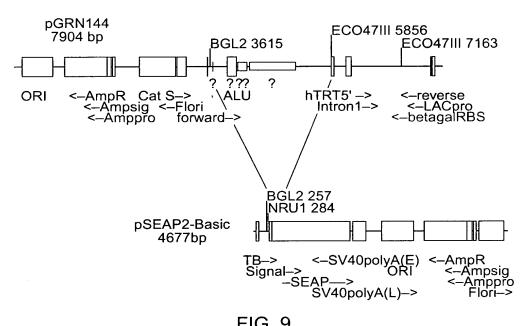


FIG. 9

+



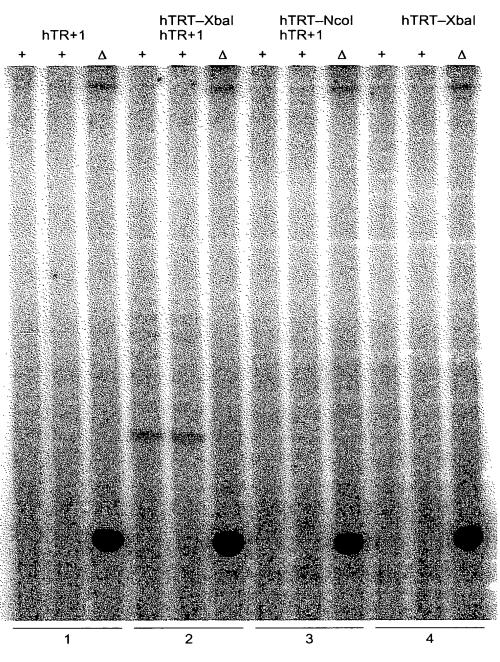


FIG. 10A

╅

+



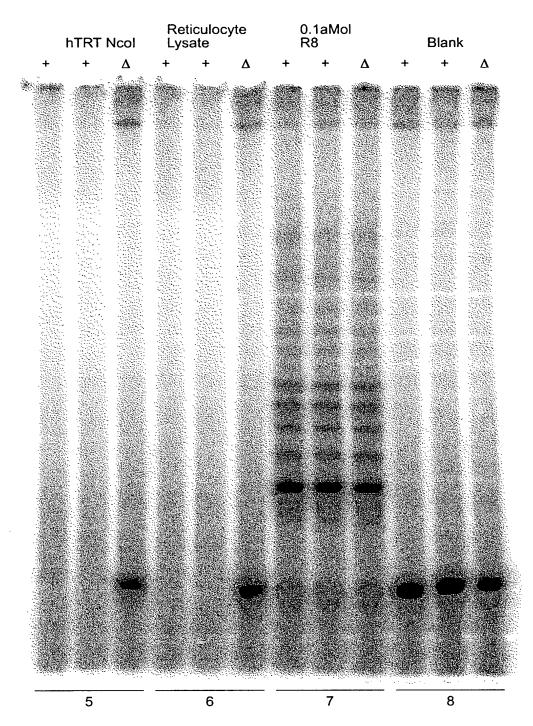


FIG. 10B



### Telomerase Specific Motifs

+

MOTIF T'	E <	EAEVR	ENNVR	EKEVE	ENNAC
		13	12	12	თ
MOTIF T	WI FFY TE Y RK W I I	546 WLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGI 13 EAEVR	429 WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFI 12 ENNVR	441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSI 12 EKEVE	366 WLFRQLIPKIIQTFFYCTEISSTVT.IVYFRHDTWNKLITPFI 9 ENNVC
		54	42	44	36
	TRT con	hTRT	$\operatorname{spTRT}$	Ea p123	Sc_Est2

## Telomerase RT Motifs (Fingers)

MOTIF A MOTIF B'	plyF D cYD i Yq GipQGslS l	PELYFVKVDVTGAYDTI 104 YVQCQGIPQGSILSTLLCSLCY	VRIDIKSCYDRI 99 YLQKVGIPQGSILSSFLCHI	67 PKLFFATMDIEKCYDSV 117 YKQTKGIPQGLCVSSILSSFYY	68 PELYFMKFDVKSCYDSI 85 YIREDGLFQGSSLSAPIVDLVY	h hDh AF h hPQG pP hh	GY
Š	p lyF	69 PELYF	66 RKKYF	67 PKLFF	68 PELYF	Ч	
MOTIF 2	fR I	0 LRPIV	O FRLIT	O FRPIM	2 FRIIA	hR h	
MOTIF 1	R iPKk	11 SRLRFIPKPDG 0 LRPIV	10 AVIRLLPKKNT 0 FRLIT	10 GKLRLIPKKTT 0 FRPIM	13 SKMRIIPKKSN 2 FRIIA	p hh h K	
	TRT con	hTRT	$\mathtt{spTRT}$	Ea_p123	Sc_Est2	RT con	

# Telomerase RT Motifs (Palm, Primer Grip)

		192	176	174	141			
MOTIF E	wgs 1	WCGLLLDTRTL	FFGFSVNMRSL	WIGISIDMKTL	WKHSSTMNNFH	hLG h		
		24	22	28	25			
		ľW	ľVI	Ę	ILA	Ч		,
Ω	n K	LRK	LEK	MKK	RDK	СK		( i
MOTIF D	g	GVPEYGCVVN	GFEKHNFSTS	VSRENGFKFN	GFQKYNAKAN	Gh h cK h		i
		15	15	15	15			
MOTIF C	lllrl DDfL it	LLLRLVDDFLLVT	16 VLLRVVDDFLFIT 15 GFEKHNFSTSLEKTVI 22 1	24 LLMRLTDDYLLIT 15 VSRENGFKFNMKKLQT 28	$\vdash$	h Y DDhhh	Œ	
		15	16	24	18			
	TRT con	hTRT	SPTRT	Ea_p123	Sc Est2	RT_con		

FIG. 11



> NFkB\_CS1 GGGRQTYYQC NFkB-MHC-I.2 TGGGCTTCCCC

### 

301 GCTGGGGTTGAGGGCGGCCGGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTCCGACCCCAACTCCCGCCGGCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

NFkB\_CS1
GGGRQTYYQC
NFkB\_CS2
RGGGRMTYYCC
Topo\_II\_cleavage\_site
RNYNNCNNGYNGKTNYNY

361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGG TCCCGCGAAGGGGGCGTCCACAGGACGACTTCCTCGACCACCGGGCTCACGACGTCTCC

FIG. 12

╫

+



1 AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT 51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA 151 TCCAGAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT 201 TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG 251 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT 301 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACTTCTT TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCA 401 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA 451 CTCAAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG 501 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA 551 GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG 601 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA 651 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA 701 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA 751 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT 801 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTC AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA 851 901 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG 951 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG 1001 1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAATC GAAAACTTGA 1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC TACACAACTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA 1151 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC 1201 AAAAGAAAGT TAAGAAATAT GTGGAACTAA ACAAGCATGA ACTCATTCAC 1251 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA GGTTGAGACC TCTGCAAAGC ATTTTTATTA TTTTGATCAC GAAAACATCT ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG 1301 1351 1401 CTGATTAGAT GATTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA 1451 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA 1501 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT 1551 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG TAAATTCAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG 1701 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC 1751 TTTTGGATTC GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG 1801 AGTTTGTTTG CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTTGCA 1851 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAACTATC 1901 AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA 1951 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC 2001 2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG 2151 2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG CCAATATAAT TACATTAACT TTAATGGGAA GTTTTATAAA CAAACAAAAG 2251 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTCATC ATTTTATTAT 2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

FIG. 13A





2401	CCCTGAAAAT	CCAAATGTTA	ATCTTCTAAT	GAGACTTACA	GATGACTATC
2451	TTTTGATTAC	AACTCAAGAG	AATAATGCAG	TATTGTTTAT	TGAGAAACTT
2501	ATAAACGTAA	GTCGTGAAAA	TGGATTTAAA	TTCAATATGA	AGAAACTACA
2551	GACTAGTTTT	CCATTAAGTC	CAAGCAAATT	TGCAAAATAC	GGAATGGATA
2601	GTGTTGAGGA	GCAAAATATT	GTTCAAGATT	ACTGCGATTG	GATTGGCATC
2651	TCAATTGATA	TGAAAACTCT	TGCTTTAATG	CCAAATATTA	ACTTGAGAAT
2701	AGAAGGAATT	CTGTGTACAC	TCAATCTAAA	CATGCAAACA	AAGAAAGCAT
2751	CAATGTGGCT	CAAGAAGAAA	CTAAAGTCGT	TTTTAATGAA	TAACATTACC
2801	CATTATTTTA	GAAAGACGAT	TACAACCGAA	GACTTTGCGA	ATAAAACTCT
2851	CAACAAGTTA	TTTATATCAG	GCGGTTACAA	ATACATGCAA	TGAGCCAAAG
2901	AATACAAGGA	CCACTTTAAG	AAGAACTTAG	CTATGAGCAG	TATGATCGAC
2951	TTAGAGGTAT	CTAAAATTAT	ATACTCTGTA	ACCAGAGCAT	TCTTTAAATA
3001	CCTTGTGTGC	AATATTAAGG	ATACAATTTT	TGGAGAGGAG	CATTATCCAG
3051	ACTTTTTCCT	TAGCACACTG	AAGCACTTTA	TTGAAATATT	CAGCACAAAA
3101	AAGTACATTT	TCAACAGAGT	TTGCATGATC	CTCAAGGCAA	AAGAAGCAAA
3151	GCTAAAAAGT	GACCAATGTC	AATCTCTAAT	TCAATATGAT	GCATAGTCGA
3201	CTATTCTAAC	TTATTTTGGA	AAGTTAATTT	TCAATTTTTG	TCTTATATAC
3251	TGGGGTTTTG	GGGTTTTGGG	GTTTTGGGG		

FIG. 13B

1	MEVDVDNQAD	NHGIHSALKT	CEEIKEAKTL	YSWIQKVIRC	RNQSQSHYKD
51	LEDIKIFAQT	NIVATPRDYN	EEDFKVIARK	EVFSTGLMIE	LIDKCLVELL
101	SSSDVSDRQK	LQCFGFQLKG	NQLAKTHLLT	ALSTQKQYFF	QDEWNQVRAM
151	IGNELFRHLY	TKYLIFQRTS	EGTLVQFCGN	NVFDHLKVND	KFDKKQKGGA
201	ADMNEPRCCS	TCKYNVKNEK	DHFLNNINVP	NWNNMKSRTR	IFYCTHFNRN
251	NQFFKKHEFV	SNKNNISAMD	RAQTIFTNIF	RFNRIRKKLK	DKVIEKIAYM
301	LEKVKDFNFN	YYLTKSCPLP	ENWRERKQKI	ENLINKTREE	KSKYYEELFS
351	YTTDNKCVTQ	FINEFFYNIL	PKDFLTGRNR	KNFQKKVKKY	VELNKHELIH
401	KNLLLEKINT	REISWMQVET	SAKHFYYFDH	ENIYVLWKLL	RWIFEDLVVS
451	LIRCFFYVTE	QQKSYSKTYY	YRKNIWDVIM	KMSIADLKKE	TLAEVQEKEV
501	EEWKKSLGFA	PGKLRLIPKK	TTFRPIMTFN	KKIVNSDRKT	TKLTTNTKLL
551	NSHLMLKTLK	NRMFKDPFGF	AVFNYDDVMK	KYEEFVCKWK	QVGQPKLFFA
601	TMDIEKCYDS	VNREKLSTFL	KTTKLLSSDF	WIMTAQILKR	KNNIVIDSKN
651	FRKKEMKDYF	RQKFQKIALE	GGQYPTLFSV	LENEQNDLNA	KKTLIVEAKQ
701	RNYFKKDNLL	QPVINICQYN	YINFNGKFYK	QTKGIPQGLC	VSSILSSFYY
751	ATLEESSLGF	LRDESMNPEN	PNVNLLMRLT	DDYLLITTQE	NNAVLFIEKL
801	INVSRENGFK	FNMKKLQTSF	PLSPSKFAKY	GMDSVEEQNI	VQDYCDWIGI
851	SIDMKTLALM	PNINLRIEGI	LCTLNLNMQT	KKASMWLKKK	LKSFLMNNIT
901	HYFRKTITTE	DFANKTLNKL	FISGGYKYMQ	CAKEYKDHFK	KNLAMSSMID
951	LEVSKIIYSV	TRAFFKYLVC	NIKDTIFGEE	HYPDFFLSTL	KHFIEIFSTK
1001	KYIFNRVCMI	LKAKEAKLKS	DQCQSLIQYD	A	

FIG. 14



1469 1018 1078 1198 1272 86 1405 128 agotottggagtagotoacagaaatoottacaaatottotgatgagactatattagattoattacagtoogtgcatatto ggtaccgatttactttcctttcttcataagctaattgcttcctcgaacgctcctaaatctctggaaatatttttacaaga actcaataacaataccaagtcaaattccaatatgaaggtgttattagtgatcgataatatttctattttatcggtcgtta gttgataattatttgcaaaatcatgtccttagtggtggtaatccgcgaaagttttttgatgcttgcacacgtctagcatg ccaaatatgtatcatctcgtattaggcttttttccgttttactcctggaatcgtacctttttcactattccccctaatga ataatctaaattagtttcgcttataattgatagtagtagaaagattggtgattctactcgtgtaatgttattagtttaaa gtatatatatttttgttttgatttttttttctattcgggatagctaatatggggcag gtaaggtattctaattgtgaaatatttacctgcaattactgtttcaaagaga ggttcgcttacttttaatcgtggtactgttttagctgctacttctagccaaccgcgtgtttctaccccgtcattggatat ttaacatggagccttacactttagatgagtcacgtcgcatgatggagtatttggtatcatccaacgtttgccttgaaaag attgagatattcaaaaaatttctatccactacaactcctttaacgcggttttatttttctattttctattctcatgttgtt GTA AAT TAT AGC CAA AAT N GTA CAA AAT N GGA GAG E AAC N TCG GAG GTA V CTA 999 CAT TTTATG ( CGC R AGT S GAA GCC A CII GAA E GAT CGA ATT GAT D CLI CCC AAA AGC AGG P K S R  $_{
m LLL}$ CAA TTC AGC ATG GAT D GTA AGA CAG TAT GAA 团 GAT ATG1406 ttgtatttaaccgataaag AAT CAT 114 N H S S S S S S S AAA ĸ Œ, AAT GAG CAC CAT ACC H H T GAA S S S S S S (H) TTA  $^{\mathrm{TCA}}$  $^{\mathrm{TGC}}$ ACC CAG ATA AAA TGI TCA CTATGC ATG CATAAA K 1019 TAC TAT 959 1079 1139 1199 1273 401 561 641 61 107 721 87 481 801

FIG. 15A



3960 818	4020 838	4089 848	4149 868	4209 888	4274 903	4339 917	4401 935	4468 946	4528 966	4588 986	4665 989
TGT	999	AG gtatactgtgtaactgaataatagctgacaaataatcag A TCG R	TCT	AAA K	G gtgagtacttattttaactaga D	3 GCC	TG gtacgtgtc W	AAA K	ACT	ATA I	tgtcattttcaatttattatacatcctttattactggtgtcttaaacaatattattactaagtata
GCA A	ATG M	ag 7	AAT N	TTA L	taa	rTG L	ytac	$_{ m L}^{ m TTG}$	$_{\rm L}^{\rm TTG}$	AGA R	taad
TTA L	CAT H	aato	TTC F	TAC Y	attt	AAG K	D. S. ™	GGT G	TCA S	AGA R	ıttac
TTG L	AAA K	aaat	AAA K	GCA A	actt	, AAA K	AAA K	GAT D	CAG Q	CAT H	atte
ACA T	ACG T	tgad	TCA	CAA	gagt	TGG 7	GTC V	AGA R	TTT F	TTA L	caat
GAT D	$^{ m CTG}_{ m L}$	tago	AAT N	GCA A	G gt D	AAA ATT 3 K I V	GAA E	ATG M	CAA O	TTT F	taaa
CTT L	GAG E	ataa	CAC H	AGA R	ACG	AAA K	GCA	GGA	TAC Y	TTA L	gtct
TCT S	GTA V	ctga	ACC CAC I	ATG M	ATA I	AGA R	TCT S	CTT L	ATA I	GTG V	tggt
AGG R	TCT	gtaa	ATT I	TGT C	TTC F	GGA	TCC	$^{ m TGT}_{ m C}$	CTA	CAG Q	ttac
ATG M	ACA T	ctgt	GAC	ATG M	ATG M	ATT I	$_{ m L}^{ m TTG}$	TTT F	CAG Q	CGA R	ttta
AAC N	TCT S	tata	ATT GAC I	TCT S	AGA R	GTT	TTC F	$_{\rm L}^{\rm CTT}$	GAA	$^{ m TTG}_{ m L}$	atco
GTG V	AAC N		TTT F	TAC Y	CAA O	AAT N	CGT R	$\tt ggtctcgagacttcagcaatattgacacatcag \ G \ CTT \\ L$	TTC F	GTT V	atac
TCT S	TTT F	CTA	GTA V	GGA G	CCC	, TTG	AGG R	atca	TGC	CCA P	ıttat
TTC F	TTA L	ATT I	CAA O	CTA L	ATT I	CTT	AGT S	acac	CCA P	AGA R	ttta
GGT G	GCC	AAA K	GCA	AGG R	TTT F	g AT	ACG	attg	CAT H	CTA L	tcaa
TTC F	GAA E	TAC Y	TTT F	TAT Y	ATA I	ctta	TAT Y	Icaat	AAA TAT ( K Y I	CCG P	attt
TTC F	GAT D	TTT F	TCC	ATA I	GAT D	taac	GGA	tcag	AAA K	AAG K	tgto
CCA P	ATT I	TTT F	GCA A	AAT N	AAG K	aaagtcattaattaaccttag	TTA L	gact	TTC F	ATC	TAA *
ATG M	AAA K	TCT S	CTT L	TGC C	ATG M	ıtcat	ATA I	tcga	TCT	CTT L	GAT D
AGA R	CCT P	AAA K	AGC S	TGC	AGG R	aaag	GAA E	ggto	CCC	GAT D	GCT A
3901 799	3961 819	4021 839	4090 849	4150 869	4210 889	4275 904	4340 918	4402 936	4469 947	4529 967	4589 987

.IG. 75

5145 5225 5305 5385 5465

5544

4825 4905 4985 5065

+



ttaaaagtaatacatgaggctaatctcctttcatttagaataaggaaagtggttttctataatgaataatgcccgcacta atgcaaaaaagacgaagattatcttctaaacaagggggattaagcatatccgaaggaaaagagagaataatacccagtgtt atgtottatataaaggttitgttitttootgacticaatittigoatgggtgaaaagaaatagtgttaagcoattattggat gttgaagaaagcaaggataattttggaacaagcttctgcagatgacaggctaaattttggtgaccgaattttggtaaaagc cccaggitatccatggitggccggccttgctactgagacgaaaagaaactaaggatagittgaatactaatagctcatita tecgaaatagecaaatttettggtteeteaaageggaagtetaaagaaetttgaagettatgaagettetgaagaetee

tratccttatacttttaagaaagattgacagtggttgctgactactgcccacatgcccattaaacgggagtggttaaaca

4746 4826

4906 4986 5066 5146 5226 5306

FIG. 15F

aaaaaagaaaatatcattgggagacatctcttgatgaatcagatgcggagagagtatctccagcggatccttgatgtcaata acttctattttctgaaatgtatggtcctactgtcgctttcgacttctcgtagctctacgcagttaagtgaccaaaggtacc

tcctgatttaaaggaggaatcttccaccgatgaggaaatggatagcttatcagctggtgaggaggaagcctaattttttgc

+

1 gcagegetge gteetgetge geaegtggga ageeetggee eeggeeacee eegegatgee gegegetece egetgeegag cegtgegete cetgetgege agecactace gegaggtget 121 gccgctggcc acgttcgtgc ggcgcctggg gccccagggc tggcggctgg tgcagcgcgg 181 ggacceggeg gettteegeg egetggtgge ceagtgeetg gtgtgegtge eetgggaege 241 aeggeegeee eeegeegeee eeteetteeg eeaggtgtee tgeetgaagg agetggtgge 301 ccgagtgctg cagaggctgt gcgagcgcgg cgcgaagaac gtgctggcct tcggcttcgc 361 gctgctggac ggggcccgcg ggggcccccc cgaggccttc accaccagcg tgcgcagcta 421 cctgcccaac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc tgctgctgcg 481 ccgcgtgggc gacgacgtgc tggttcacct gctggcacgc tgcgcgctct ttgtgctggt 541 ggctcccagc tgcgcctacc aggtgtgcgg gccgccgctg taccagctcg gcgctgcac 601 tcaggcccgg cccccgccac acgctagtgg accccgaagg cgtctgggat gcgaacgggc 661 ctggaaccat agcgtcaggg aggccggggt ccccctgggc ctgccagccc cgggtgcgag qaqqqqqqq ggcagtgcca gccgaagtct gccgttgccc aagaggccca ggcgtggcgc 781 tgcccctgag ccggagcgga cgcccgttgg gcaggggtcc tgggcccacc cgggcaggac 841 gegtggaceg agtgacegtg gtttetgtgt ggtgteacet geeagaceeg eegaagaage 901 cacetettg gagggtgege tetetggeae gegeeactee cacecateeg tgggeegeea 961 gcaccacgcg ggccccccat ccacatcgcg gccaccacgt ccctgggaca cgccttgtcc 1021 cccggtgtac gccgagacca agcacttcct ctactcctca ggcgacaagg agcagctgcg 1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga ggctcgtgga 1141 gaccatettt etgggtteca ggeeetggat geeagggaet eeeegeaggt tgeeeegeet 1201 gccccagcgc tactggcaaa tgcggcccct gtttctggag ctgcttggga accacgcgca 1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgctg cgagctgcgg tcaccccagc 1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggcccccg aggaggagga 1381 cacagaccc cgtcgcctgg tgcagctgct ccgccagcac agcagcccct ggcaggtgta 1441 cggcttcgtg cgggcctgcc tgcgccggct ggtgccccca ggcctctggg gctccaggca 1501 caacgaacgo cgottootoa ggaacaccaa gaagttoato toootgggga agcatgooaa 1561 gctctcgctg caggagctga cgtggaagat gagcgtgcgg gactgcgctt ggctgcgcag 1621 gagcccaggg gttggctgtg ttccggccgc agagcaccgt ctgcgtgagg agatcctggc 1681 caagtteetg caetggetga tgagtgtgta egtegtegag etgeteaggt etttettta 1741 tgteaeggag accaegttte aaaagaacag getetttte taeeggaaga gtgtetggag 1801 caagttgcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc gggagctgtc 1861 ggaagcagag gtcaggcagc atcgggaagc caggcccgcc ctgctgacgt ccagactccg 1921 cttcatccc aagcctgacg ggctgcggcc gattgtgaac atggactacg tcgtgggagc 1981 cagaacgttc cgcagagaaa agagggccga gcgtctcacc tcgagggtga aggcactgtt 2041 cagegtőete aactaegage gggegeggeg ceceggeete etgggegeet etgtgetggg 2101 cctggacgat atccacaggg cctggcgcac cttcgtgctg cgtgtgcggg cccaggaccc 2161 gccgcctgag ctgtactttg tcaaggtgga tgtgacgggc gcgtacgaca ccatcccca ggacaggete acggaggtea tegecageat cateaaacee cagaacaegt actgegtgeg 2281 teggtatgee gtggteeaga aggeegeeca tgggeaegte egeaaggeet teaagageea 2341 cgtctctacc ttgacagacc tccagccgta catgcgacag ttcgtggctc acctgcagga 2401 gaccagccg ctgagggatg ccgtcgtcat cgagcagagc tcctccctga atgaggccag 2461 cagtggcete ttegaegtet teetaegett catgtgeeae caegeegtge geateagggg 2521 caagteetae gteeagtgee aggggateee geagggetee atecteteea egetgetetg 2581 cageetgtge taeggegaea tggagaacaa getgtttgeg gggattegge gggaeggget gctcctgcgt ttggtggatg atttcttgtt ggtgacacct cacctcaccc acgcgaaaac 2641 cttcctcagg accctggtcc gaggtgtccc tgagtatggc tgcgtggtga acttgcggaa 2761 gacagtggtg aacttccctg tagaagacga ggccctgggt ggcacggctt ttgttcagat 2821 gccggcccac ggcctattcc cctggtgcgg cctgctgctg gatacccgga ccctggaggt 2881 gcagagcgac tactccagct atgcccggac ctccatcaga gccagtctca ccttcaaccg 2941 cggcttcaag gctgggagga acatgcgtcg caaactcttt gggggtcttgc ggctgaagtg 3001 teacageetg tttetggatt tgeaggtgaa cageeteeag aeggtgtgea ceaacateta 3061 caagateete etgetgeagg egtacaggtt teacgeatgt gtgetgeage teccatttea 3121 tcagcaagtt tggaagaacc ccacattttt cctgcgcgtc atctctgaca cggcctccct 3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg ctgggggcca agggcgccgc 3241 cggccctctg ccctccgagg ccgtgcagtg gctgtgccac caagcattcc tgctcaagct 3301 gactcgacac cgtgtcacct acgtgccact cctggggtca ctcaggacag cccagacgca getgagtegg aageteeegg ggaegaeget gaetgeeetg gaggeegeag ceaaceegge 3421 actgecetea gaetteaaga ceateetgga etgatggeea eeegeeeaca geeaggeega 3481 gagcagacac cagcagccct gtcacgccgg gctctacgtc ccagggaggg aggggcggcc cacacceagg ecogeacege tgggagtetg aggeetgagt gagtgtttgg ecgaggeetg 3601 catgtccggc tgaaggetga gtgtccggct gaggectgag cgagtgtcca gccaagggct 3661 gagigtecag cacacetgee gretteaett ceccacagge tggegetegg etecaceca gggccagett tteeteacea ggageeegge tteeacteee cacataggaa tagteeatee ccagattege cattgtteac ecetegeect geeeteettt geetteeace eceaecatee 3841 aggtggagac cctgagaagg accctgggag ctctgggaat ttggagtgac caaaggtgtg 3901 cctgtacac aggcgaggac cctgcacctg gatgggggtc cctgtgggtc aaattggggg 3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa aaaaa

FIG. 16

+

MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPOGWRLVORGDP AAFRALVAQCLVCVPWDARPPPAAPSFRQVSCLKELVARVLQRL CERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPNTVTDALR GSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY OLGAATOARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG ARRRGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRG PSDRGFCVVSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPP STSRPPRPWDTPCPPVYAETKHFLYSSGDKEQLRPSFLLSSLRP SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLEL LGNHAOCPYGVLLKTHCPLRAAVTPAAGVCAREKPQGSVAAPEE EDTDPRRLVQLLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNE RRFLRNTKKFISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC VPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTETTFQKNR LFFYRKSVWSKLOSIGIROHLKRVOLRELSEAEVROHREARPAL LTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKA LFSVLNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPP ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ KAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI EOSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPQGSI LSTLLCSLCYGDMENKLFAGIRRDGLLLRLVDDFLLVTPHLTHA KTFLRTLVRGVPEYGCVVNLRKTVVNFPVEDEALGGTAFVOMPA HGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR NMRRKLFGVLRLKCHSLFLDLQVNSLQTVCTNIYKILLLQAYRF HACVLQLPFHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGSLRTAQ TQLSRKLPGTTLTALEAAANPALPSDFKTILD

### FIG. 17

TTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTG GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGG AGCCAGAACGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACT GTTCAGCGTGCTCAACTACGAGCGGGCGCGCGCCCCGGCCTCCTGGGCGCCTCTGTGCT GGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCCAGGA CCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGCGCGTACGACACCATCCC CCAGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAAACCCCAGAACACGTACTGCGT GCGTCGGTATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAG CCACGTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCCACGCTGCTCT GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGGATTCGGCGGGACGGGC TGCTCCTGCGTTTGGTGATGATTTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAA CCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACTTGCGGA AGACAGTGGTGAACTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGA TGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCTGCTGCATACCCGGACCCTGGAGG TGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC GCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGT GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGTGTGCACCAACATCT ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTC ATCAGCAAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCATCTCTGACACGGCCTCCC TCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGCCCAAGGGCCCG CCGGCC7TCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGC TGACTCGACACCGTGTCACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGC AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGAGCCAACCCGG CACTGCCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCG GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGC TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC AGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC CCCAGATTCGCCATTGTTCACCCCTCGCCCTGCCCTCCTTTGCCTTCCACCCCCACCATC CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGG AAAAAAAAAAAA

+

MetSerValTyrValValGluLeuLeuArgSerPhePhe TyrValThrGluThrThrPheGlnLysAsnArgLeuPhe PheTyrArgLysSerValTrpSerLysLeuGlnSerIle GlyIleArgGlnHisLeuLysArgValGlnLeuArgGlu LeuSerGluAlaGluValArgGlnHisArgGluAlaArg ProAlaLeuLeuThrSerArgLeuArgPheIleProLys ProAspGlyLeuArgProIleValAsnMetAspTyrVal ValGlyAlaArgThrPheArgArgGluLysArgAlaGlu ArgLeuThrSerArgValLysAlaLeuPheSerValLeu AsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAla SerValLeuGlyLeuAspAspIleHisArgAlaTrpArg ThrPheValLeuArgValArgAlaGlnAspProProPro GluLeuTyrPheValLysValAspValThrGlyAlaTyr AspThrIleProGlnAspArgLeuThrGluValIleAla SerIleIleLysProGlnAsnThrTyrCysValArgArg TyrAlaValValGlnLysAlaAlaHisGlyHisValArg LysAlaPheLysSerHisValLeuArgProValProGly AspProAlaGlyLeuHisProLeuHisAlaAlaLeuGln ProValLeuArgArgHisGlyGluGlnAlaValCysGly AspSerAlaGlyArgAlaAlaProAlaPheGlyGly

### FIG. 19

GCAGCGCTGCGT	rcctgctgcgc	\CGTGGGAAG(	CCCTGGCCCCGG	1 me GCCACCCCCGCG AT	t
pro arg ala CCG CGC GCT	pro arg cys	s arg ala v	10 val arg ser GTG CGC TCC	leu leu arg se CTG CTG CGC AG	r
his tyr arg	20 glu val let GAG GTG CTG	pro leu a	ala thr phe GCC ACG TTC	30 val arg arg le GTG CGG CGC CT	u G
				asp pro ala al GAC CCG GCG GC	
phe arg ala TTC CGC GCG	50 leu val ala CTG GTG GCG	a gln cys I C CAG TGC (	leu val cys CTG GTG TGC	60 val pro trp as GTG CCC TGG GA	p .C
ala arg pro GCA CGG CCG	pro pro ala	a ala pro s G GCC CCC T	70 ser phe arg TCC TTC CGC	gln val ser cy CAG GTG TCC TG	s
leu lys glu CTG AAG GAG	80 leu val ala CTG GTG GCG	a arg val 1 C CGA GTG (	leu gln arg CTG CAG AGG	90 leu cys glu ar CTG TGC GAG CG	g
gly ala lys GGC GCG AAG	asn val le	ı ala phe	100 gly phe ala GGC TTC GCG	leu leu asp gl CTG CTG GAC GG	y G
ala arg gly GCC CGC GGG	gly pro pro GGC CCC CCC	o glu ala p C GAG GCC :	phe thr thr TTC ACC ACC	ser val arg se AGC GTG CGC AG	r

FIG. 20A

tyr TAC	leu CTG	pro CCC	asn AAC	thr ACG	val GTG	thr ACC	asp GAC	130 ala GCA	leu CTG	arg CGG	gly GGG	ser AGC	gly GGG	ala GCG
			140 leu CTG											
			arg CGC											
ala GCC	tyr TAC	gln CAG	170 val GTG	cys TGC	gly GGG	pro CCG	pro CCG	leu CTG	tyr TAC	gln CAG	leu CTC	gly GGC	180 ala GCT	ala GCC
			arg CGG											
			200 glu GAA											
val GTC	pro CCC	leu CTG	gly GGC	leu CTG	pro CCA	ala GCC	pro CCG	220 gly GGT	ala GCG	arg AGG	arg AGG	arg CGC	gly GGG	gly GGC
			230 arg CGA											
			glu GAG											
ala GCC	his CAC	pro CCG	260 gly GGC	arg AGG	thr ACG	arg CGT	gly GGA	pro CCG	ser AGT	asp GAC	arg CGT	gly GGT	270 phe TTC	cys TGT
			pro CCT											
gly GGT	ala GCG	leu CTC	290 ser TCT	gly	thr ACG	arg CGC	his CAC	ser TCC	his CAC	pro CCA	ser TCC	val GTG	300 gly GGC	arg CGC
gln CAG	his CAC	his CAC	ala GCG	gly GGC	pro CCC	pro CCA	ser TCC	310 thr ACA	ser TCG	arg CGG	pro CCA	pro CCA	arg CGT	pro CCC
trp TGG	asp GAC	thr ACG	320 pro CCT	cys TGT	pro CCC	pro CCG	val GTG	tyr TAC	ala GCC	glu GAG	thr ACC	lys AAG	330 his CAC	phe TTC

FIG. 20B

+

							_							• • • •
leu CTC	tyr TAC	ser TCC	ser TCA	gly GGC	asp GAC	lys AAG	qlu	gln CAG	leu CTG	arg CGG	pro CCC	ser TCC	phe TTC	leu CTA
leu CTC	ser AGC	ser TCT	350 leu CTG	arg AGG	pro CCC	ser AGC	leu CTG	thr ACT	gly GGC	ala GCT	arg CGG	arg AGG	360 leu CTC	val GTG
glu GAG	thr ACC	ile ATC	phe TTT	leu CTG	gly GGT	ser TCC	arg AGG	370 pro CCC	trp TGG	met ATG	pro CCA	gly GGG	thr ACT	pro CCC
arg CGC	arg AGG	leu TTG	380 pro CCC	arg CGC	leu CTG	pro CCC	gln CAG	arg CGC	tyr TAC	trp TGG	gln CAA	met ATG	390 arg CGG	pro CCC
leu CTG	phe TTT	leu CTG	glu GAG	leu CTG	leu CTT	gly GGG	asn AAC	400 his CAC	ala GCG	gln CAG	cys TGC	pro CCC	tyr TAC	gly GGG
val GTG	leu CTC	leu CTC	410 lys AAG	thr ACG	his CAC	cys TGC	pro CCG	leu CTG	arg CGA	ala GCT	ala GCG	val GTC	420 thr ACC	pro CCA
ala GCA	ala GCC	gly GGT	val GTC	cys TGT	ala GCC	arg CGG	glu GAG	430 lys AAG	pro CCC	gln CAG	gly GGC	ser TCT	val GTG	ala GCG
ala GCC	pro CCC	glu GAG	440 glu GAG	glu GAG	asp GAC	thr ACA	asp GAC	pro CCC	arg CGT	arg CGC	leu CTG	val GTG	450 gln CAG	leu CTG
					ser AGC									
ala GCC	cys TGC	leu CTG	470 arg CGC	arg CGG	leu CTG	val GTG	pro CCC	pro CCA	gly GGC	leu CTC	trp TGG	gly GGC	480 ser TCC	arg AGG
his CAC	asn AAC	glu GAA	arg CGC	arg CGC	phe TTC	leu CTC	arg AGG	490 asn AAC	thr ACC	lys AAG	lys AAG	phe TTC	ile ATC	ser TCC
leu CTG	gly GGG	lys AAG	500 his CAT	ala GCC	lys AAG	leu CTC	ser TCG	leu CTG	gln CAG	glu GAG	leu CTG	thr ACG	510 trp TGG	lys AAG
met ATG	ser AGC	val GTG	arg CGG	asp GAC	cys TGC	ala GCT	trp TGG	520 leu CTG	arg CGC	arg AGG	ser AGC	pro CCA	gly GGG	val GTT
gly GGC	cys TGT	val GTT	530 pro CCG	ala GCC	ala GCA	glu GAG	his CAC	arg CGT	leu CTG	arg CGT	glu GAG	glu GAG	540 ile ATC	leu CTG

FIG. 20C

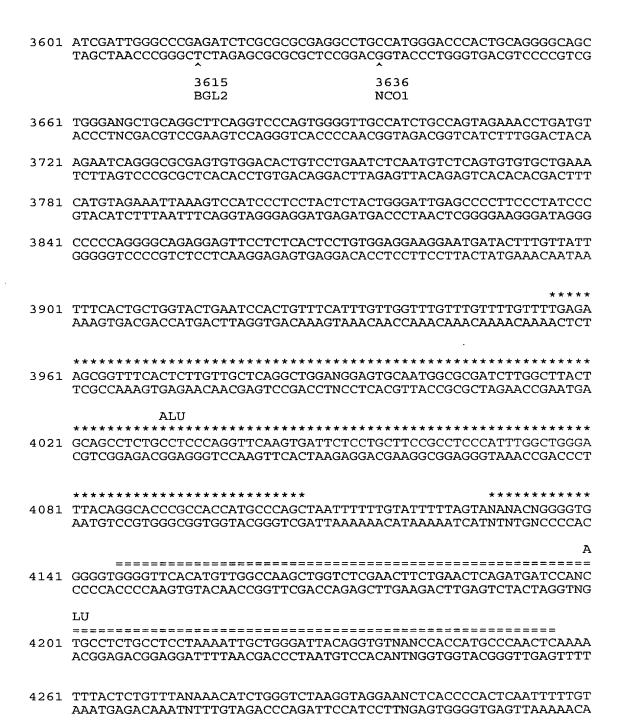
								550						
					trp TGG		met	ser						
leu CTC	arg AGG	ser TCT	560 phe TTC	phe TTT	tyr TAT	val GTC	thr ACG	glu GAG	thr ACC	thr ACG	phe TTT	gln CAA	570 lys AAG	asn AAC
arg AGG	leu CTC	phe TTT	phe TTC 590	tyr TAC	arg CGG	lys AAG	ser AGT	580 val GTC	trp TGG	ser AGC	lys AAG	leu TTG	gln CAA 600	ser AGC
ile ATT	gly GGA	ile ATC	arg AGA	gln CAG	his CAC	leu TTG	lys AAG	arg AGG	val GTG	gln CAG	leu CTG	arg CGG	glu GAG	leu CTG
					arg AGG									
leu CTG	thr ACG	ser TCC	620 arg AGA	leu CTC	arg CGC	phe TTC	ile ATC	pro CCC	lys AAG	pro CCT	asp GAC	gly GGG	630 leu CTG	arg CGG
pro CCG	ile ATT	val GTG	asn AAC	met ATG	asp GAC	tyr TAC	val GTC	640 val GTG	gly GGA	ala GCC	arg AGA	thr ACG	phe TTC	arg CGC
arg AGA	glu GAA	lys AAG	650 arg AGG	ala GCC	glu GAG	arg CGT	leu CTC	thr ACC	ser TCG	arg AGG	val GTG	lys AAG	660 ala GCA	leu CTG
phe TTC	ser AGC	val GTG	leu CTC	asn AAC	tyr TAC	glu GAG	arg CGG	670 ala GCG	arg CGG	arg CGC	pro CCC	gly GGC	leu CTC	leu CTG
gly GGC	ala GCC	ser TCT	680 val GTG	leu CTG	gly GGC	leu CTG	asp GAC	asp GAT	ile ATC	his CAC	arg AGG	ala GCC	690 trp TGG	arg CGC
thr ACC	phe TTC	val GTG	leu CTG	arg CGT	val GTG	arg CGG	ala GCC	700 gln CAG	asp GAC	pro CCG	pro CCG	pro CCT	glu GAG	leu CTG
					asp GAT									pro CCC
gln CAG	asp GAC	arg AGG	leu CTC	thr ACG	glu GAG	val GTC	ile ATC	730 ala GCC	ser AGC	ile ATC	ile ATC	lys AAA	pro CCC	gln CAG
asn AAC	thr ACG	tyr TAC	740 cys TGC	val GTG	arg CGT	arg CGG	tyr TAT	ala GCC	val GTG	val GTC	gln CAG	lys AAG	750 ala GCC	ala GCC

FIG. 20D

760 his gly his val arg lys ala phe lys ser his val leu arg pro CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA val pro gly asp pro ala gly leu his pro leu his ala ala leu GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG 790 gln pro val leu arg arg his gly glu gln ala val cys gly asp CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT ser ala gly arg ala ala pro ala phe gly gly OP TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTTGGT GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA GTATGGCTGCGTGAACTTGCGGAAGACAGTGGTGAACTTCCCTGTAGAAGACGAGGC CCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCT GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTC CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA ACTCTTTGGGGTCTTGCGGCTGAAGTGTCACAGCCTGTTTCTGGATTTGCAGGTGAACAG CCTCCAGACGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA CGCATGTGTGCTGCAGCTCCCATTTCATCAGCAAGTTTGGAAGAACCCCACATTTTTCCT GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG GATGTCGCTGGGGGCCAAGGGCGCCGCCGCCTCTGCCCTCCGAGGCCGTGCAGTGGCT GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACGTGCCACTCCT GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG ATGGCCACCGCCCACAGCCAGGCCGAGAGCAGACCAGCAGCCCTGTCACGCCGGGCT CTACGTCCCAGGGAGGGGGGGCCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG GCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC CACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTC CACTCCCCACATAGGAATAGTCCATCCCCAGATTCGCCATTGTTCACCCCTCGCCCTGCC CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT GGGGGTCCCTGTGGGTCAAATTGGGGGGGGGGTGCTGTGGGAGTAAAATACTGAATATATG 

FIG. 20E

+



**FIG. 21A** 

+

5041 NGCCANGRAGGGGCCAGGTTCCAANTTCCCAACCKTTTTWGGARGGACNGCCCCCAGGG NCGGTNCYTCCCCGGTCCAAGGTTNAAGGGTTGGMAAAAWCCTYCCTGNCGGGGGTCCC 5101 GGGGATRAACAGANTNGGGGGKGGTWGGGTTNAKGGTGGGAACNCCTTNGCGCCTGGAG CCCCTAYTTGTCTNANCCCCCMCCAWCCCAANTMCCACCCTTGNGGAANCGSCGGACCTC 5161 AACGTGCAAAGAGGAAATGAAGGGCCTGKGTCAAGGAGCCCAAGTNGGCGGGGRAGTTTG TTGCACGTTTCTCCTTTACTTCCCGGACMCAGTTCCTCGGGTTCANCCGCCCCYTCAAAC 5221 CAGGGAGGCACTCCGGGGAGGTCCSGCGTGCCCGTCCAAGGGAGCAATGCGTCCTTCGGG GTCCCTCCGTGAGGCCCCTCCAGGSCGCACGGGCAGGTTCCCTCGTTACGCAGGAAGCCC 5281 TTCGTCCCCAWGCCGCGTCTACGCGCCTYCCGTCCTCCCCTTCACGTTCCGGCATTCGTG AAGCAGGGGTWCGGCGCAGATGCGCGGARGGCAGGAGGGGAAGTGCAAGGCCGTAAGCAC 5341 GTGCCCGGAGCCCGACGCCCCGCGTCCGGACCTGGAGCCCCTGGGTCTCCGGATCAG CACGGGCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTCGGGACCCAGAGGCCTAGTC 5401 GCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCCAGGGCCTCCACATCATGGCCCCT CGGTCGCCGGTTTCCCAGCGGCGTGCGTGGACAAGGGTCCCGGAGGTGTAGTACCGGGGA

5461 CCCTCGGGTTACCCCACAGCCTAGGCCGGATTCGACCTCTCTCCGCTGGGGCCCTCGCCT GGGAGCCCAATGGGGTGTCGGATCCGGCCTAAGCTGGAGAGAGGCGACCCCGGGAGCGGA Sp1 5521 GGCGTCCTGCACCTGGGAGCGCGAGCGCGCGCGGGGGGGAAGCGCGGCCCATACCC CCGCAGGGACGTGGGACCCTCGCGCTCGCCGCGCCCCCCTTCGCGCCCGGGTATGGG 5581 CCGGGTCCGCCGGAAGCAGCTGCGCTGTCGGGGCCAGGCCGGGCTCCCAGTGGATTCGC GGCCCAGGCGGGCCTTCGTCGACGCGACAGCCCCGGTCCGGCCCGAGGGTCACCTAAGCG Topo II cleavage site \*\*\*\*\*\* 5641 GGGCACAGACGCCCAGGACCGCGCTTCCCACGTGGCGGAAGGACTGGGGACCCGGGCACC CCCGTGTCTGCGGGTCCTGGCGCGAAGGGTGCACCGCCTTCCTGACCCCTGGGCCCGTGG E2F \*\*\*\*\* 5701 CGTCCTGCCCTTCACCTTCCAGCTCCGCTTCTTCCGCGCGGGCCCCGGCCCCGTCCCGAA GCAGGACGGGGAAGTGGAAGGTCGAGGCGAAGAAGGCGCCCTGGGCCGGGCCAGGGCTT Ε Sp1 ======= NFkB h 2F \*\*\*\*\* \*\*\*\* 5821 CGCGGCCCGCCTCTCCTTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCCTGCTGCGCA GCGCCGGGCGGAGAGGAAGCGCCGCTCAAAGTCCGTCGCGACGCAGGACGACGCGT 5860 5875 ECO47III FSP1 TRT5' \*\*\*\*\*\*\*\* 5881 CGTGGGAAGCCCTGGCCCCGGCCACCCCCGCGATGCCGCGCTCCCCGCTGCCGAGCCG GCACCCTTCGGGACCGGGGCCGGTGGGGGGCGCTACGGCGCGCGAGGGGCGACGGCTCGGC 5941 TGCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTCGTGCGGC ACGCGAGGGACGACGCGTCGGTGATGGCGCTCCACGACGCGACCGGTGCAAGCACGCCG 5953 FSP1 6001 GCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACCCGGCGGCTTTCCGCGCGC CGGACCCGGGGTCCCGACCGCCGACCACGTCGCGCCCCTGGGCCGCCGAAAGGCGCGCG NFkB

FIG. 21C

\_\_\_\_\_

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* 6121 CCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGG GGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGCCGACCCCAACTCCCGCCGGCCCC Topo\_II\_cleavage\_s NFkB Intron1 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* 6181 GGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTCAGGGCGCTTCCCCCGCAGGTGTC CCTTGGTCGCTGTACGCCTCTCGTCGCGTCGCTGAGTCCCGCGAAGGGGGGCGTCCACAG ite GACGGACTTCCTCGACCACCGGGCTCACGACGTCTCCGACACGCTCGCGCCGCGCTTCTT 6301 CGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCGCGGGGGCCCCCCGAGGCCTT GCACGACCGGAAGCCGACGACCTGCCCCGGGCGCCCCCGGGGGGGCTCCGGAA 6361 CACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCGGGGGGAGCGG GTGGTGGTCGCACGCGTCGATGGACGGGTTGTGCCACTGGCTGCGTGACGCCCCCTCGCC 6372 FSP1 6421 GGCGTGGGGGCTGCTGCGCGCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGCACG CCGCACCCCGACGACGCGCGCGCGCCCCGCTGCTGCACGACCAAGTGGACGACCGTGC 6481 CTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGCCGCCGCT GACGCGGGGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCGGCGGCGA 6541 GTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCCGAAG CATGGTCGAGCCGCGACGGTGAGTCCGGGCCGGGGCGGTGTGCGATCACCTGGGGCTTC 6601 GCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCTGGG CGCAGACCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCC 6661 CCTGCCAGCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCC GGACGGTCGGGGCCCACGCTCCTCCGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGG 6721 CAAGAGGCCCAGGCGTGCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTC GACCCGGGTGGGCCCGTCCTGCGCACCTGGCTCACTGGCACCAAAGACACACCACAGTGG 6841 TGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTC ACGGTCTGGGCGCTTCTTCGGTGGAGAAACCTCCCACGCGAGAGACCGTGCGCGGTGAG 6901 CCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCATCCACATCGCGGCCACCACG GGTGGGTAGGCACCCGGCGGTCGTGGTGCGCCCGGGGGGTAGGTGTAGCGCCGGTGGTGC

+

6961	TCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC AGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAG
7021	${\tt AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCTGACTCGCTGTTCCTCGTCGACGCCGGGAGGAAGGA$
7081	${\tt TGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCCCTGGATGCCAGGGACACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGAACGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGTAGAAAGACCCAAGGTCCCAAGGTCCGGGACCTACGGTCCCTGGAACGTCCCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTGGAACGTCCTGGAACGTCCTGGAACGTCCTGGAACGTCCTGGAACGTCCCTGGAAGGTCCTACGGTCCCTGGAACGTCCTGGAACGTCCTGAAGAAGACCCAAGGTCCTACGGTCCCTGGAACGTCCTACGGTCCCTGGAACGTCCTGAAGAAGACCCAAGGTCCTACGGTCCTGGAACGTCCTACGGTCCTGGAACGTCCTACGGTCCTGGAACGTCCTACGGTCCTGAAAGAAGACCCAAGGTCCTACGGTCCTGAAAGAAGACCCAAGGTCCTACGGTCCTGAAAGAAGAACAAAGACCCAAGAAGAACAAAGAACAAAAGAACAAAAAA$
7141	${\tt TCCCCGCAGGTTGCCCCGCCTGCCCAGCGCTACTGGCAAATGCGGCCCCTGTTTCTGGAAGGGGCGTCCAACGGGGCGGACGAGGGTCGCGATGACCGTTTACGCCGGGGACAAGACCT}$
	7167 ECO47III
7201	${\tt GCTGCTTGGGAACCACGCGCAGTGCCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCTCGACGAACCCTTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGGCGACGAGGAGTTCTGCGTGACGGGCGACGAGGAGTTCTGCGTGACGGGCGACGAGGAGTTCTGCGTGACGGGCGACGAGGAGTTCTGCGTGACGGGCGAACCCTTGCGTGACGGGCGAACCCCTACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCGCGAGGAGTTCTGCCGCGAACCGAGGAGTTCTGCGTGACGGGCGAACCCCACGAGGAGTTCTGCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACGAGGAGTTCTGCCCCACAGAGAGAG$
7261	${\tt GCGAGCTGCGGTCACCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGTCGCTCGACGCCAGTGGGGTCGTCGGCCACAGACACGGGCCCTCTTCGGGGTCCCGAGACACACAC$
7321	${\tt GGCGGCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCTCCGCCAGCACGCCGGGGGGGCTCCTCCTCTGTGTCTGGGGGCAGCGACCACGTCGACGAGGGGGTCGT}$
7381	${\tt CAGCAGCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCT$
7441	${\tt AGGCCTCTGGGGCTCCAGGCACCAACGAACGCCGCTTCCTCAGGAACACCCAAGAAGTTCATTCCGGAGACCCCGAGGTCCGTGTTGCTTGC$
7501	$\tt CTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCCTCGACTGCACCTTCTACTCGCACGC$
	**********
7561	GGACTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCCAGG CCTGACGCGAACCGACGCGTCCTCGGGTCCACCACCACCGGCAGCTCCCGGGTCC
	7575 FSP1
	Intron2
	************
7621	$\tt CCCCAGAGCTGAATGCAGTAGGGGGCTCAGAAAAGGGGGCAGGCA$
	**************
7681	GTCTCCATCGTCACGTGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACGGTGCAGAGGTAGCAGTGCACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGCCAC
	**>
7741	ATCGAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATA TAGCTCCAGCTGAGATCTCCTAGGGGCCCCATGGCTCGAGCTTAAGCATTAGTACCAGTAT
	7747

FIG. 21E

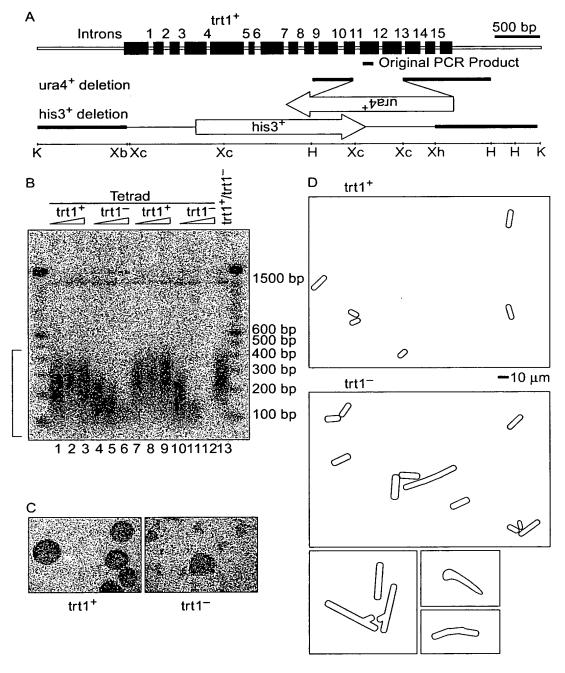
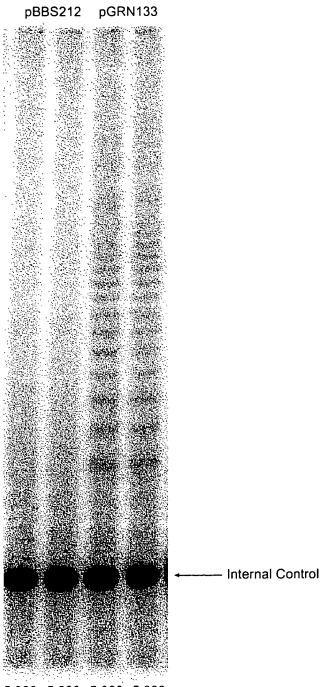


FIG. 22

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG ACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC AGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATC AGGGGCAAGTC

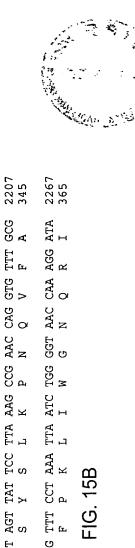
FIG. 24



Approximate Cell No. 5,000 5,000 5,000 5,000

FIG. 25

+



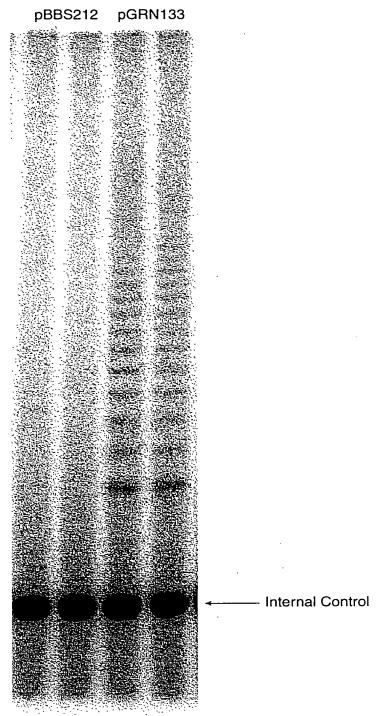
1601 155 1841 235 1907 245 1967 265 2027 285 2087 305 2147 325 1781 215 caa o GTG V GAC D TAT Y AGG R TTA ATT GAA CAA ACA L I E Q T ATT GCG A AAT AAA AGC GCC CGC AAA N K S A R K GTG ATT CCA CTG GTA V I P L V AAC N T ATC I TCC S TAT Y TGG W AAT N ACT CCA P TTT F CCA P CCA GAA AAA AAT K N TTT F TAC AGG TCA Y R S gtaaataccggttaagatgttgcgcactttgaacaagactgacaagtatag TAT Y TAT TGC ( ATT I  $_{\rm L}^{\rm CTT}$ CTA L TGG M AAC N TCA S GCT A ΑT CAC AAA H K GAG E GAG E GTG V TTT F CCT P CAT H CCG P CAA O gtaactaatactgttatccttcataactaattttag CAA TAC AAC AAT N  $_{\rm L}^{\rm CTT}$ TTT F ATT I CTT L AAG K ATA AAT N TTG TAC GTA V ACT T AGC S ATT I TGG W TCC ATT S I AAG K AAA K TTT F ATG M TCT S AGT S GTG AAG CAA V K Q GTT V TCC S CTA L AAA K ATA I TTT F AGG R 66**3**4 CAC H TAT Y CTC TCA TAC CTT Y L ACA T CTT L AGT S GTA V AGT S AAA K ATT TCT GGC ATA CCA I S G I P GAA E ATT I CAA O CTA L TCC S ACA T GTG CCC AAA CGT V P K R  $_{\rm L}^{\rm CTT}$ AAT N TTA L AAG CGA ACC ATT K R T I TCA S AAC N ATA AAC GCA TTT I N A F TCT S CGA R AGC ATT S CCT CGT ATT R I AAA K GTT V CGG R GAA E GAT D യ മ CAT H AAT N GTT V CAA O CAT H TGT C GAT D AT GAA E ATG M CAG AGAR TGG W CTT L ACA T CTC L AAG K ATT I GAT D TCT S GCC A GGA G AGT S CGA R CTT L AAA K TCC TCT S TTT F TCA CAG A AAG K TAC Y CAC H GAT D AAA K GTT V AAG K TTT F AAT N TCA S AAG K TTA L CAA Q GCA A AGT S GAA E  $_{\rm L}^{\rm CTT}$ 1602 156 1662 176 1782 216 1842 236 1908 246 1968 266 2028 286 2088 306 1722 196

2465 405 2525 425 2585 445 2645 465 2705 485 2775 495 2835 515 2906 524 2967 542 3027 562 3088 581 gtaat TTC TTA ATA AAG gtattaatttttggtcatcaatgtactttacttctaatctatta F  $\,$  L  $\,$  I  $\,$  K 2706 gtattttaaagtatttttgcaaaagctaattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 486 CTC GAA ACT L E T GAG E CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG R L L P K K N T F R L I T GTG V GAG E TTG L TGG W AAC N CTT CCT P TTA CGA ( ATA AAC N GTC V AAA G gtattgtataaaatttattaccactaacgattttaccag AC K D AAA K GAT D TTT F AGT S CAA ACT GAA E TGG M CCA P AAG K AAA K CAA TTA L GAA CGC R AGA R  $ext{TTT}$ ATT TTA L ATG TTA GTC AGT ACG AAC M L V S T N AAG K TAT Y GAG AAA E K TTT F GCG A GGT G TCA GAA ATT S E I ATT I AGT AGT S S CAT H CCT P TAT Y GAA E CTT L ATA ATA I I TTT F GTT V CTT L TCA ATG AAA ATG S M K M GAA GAA I GAT CGA ACT R T GAT D AGT S 2397 gtaatatgccaaatttttttaccattaattaacaatcag ATT 396 TTA AGT ( L S I TTT AAG AAG K K AAT N AAT N TCG AGA TAC GAG TCT S R Y E S ATT ACA I T TGC C CGA R AAA K ATC I AAT N AAA ATG 1 K M C GTT ATT (V I I TAC Y TTA Ľ AAA K TTA L AAC N CAT H CTA L GAT D  $ext{TTT}$ GCG A CCA GCA P A AGA R TTA TGG W AGT S င္လင္လင္လ မ TCA S AAA K ATA I AAT TAC TTG TGC CGA L C R AAA K TCA S GGT G CTG L TTA AGA A GAG ATA I TTG AAA L K TCA S ATC I TTG CCT L P ATG M ATA I GAA E AGG R TTC ACT T TCG S 2907 ttagcag 525 2466 AAA . 406 K 2646 CTC 7 466 L 2776 ACT 496 T 2836 AAT ' 516 N 2586 ATC 446 I 2526 GAA 426 E  $_{
m LLL}$ TTC F GCA A GTT V 3028 366

FIG. 15C



<u>-1G. 151</u>



Approximate Cell No. 5,000 5,000 5,000 5,000

FIG. 25

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG ACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC AGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATC AGGGGCAAGTC

FIG. 24

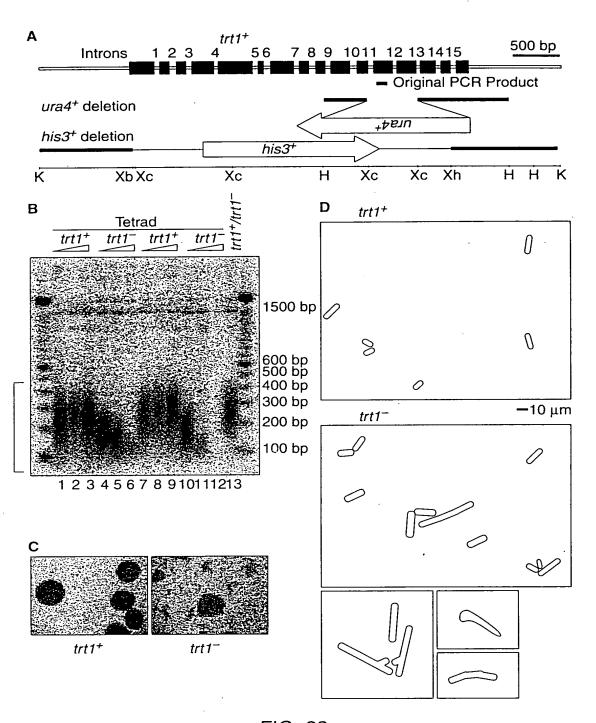


FIG. 22

6961	TCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC AGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAG
7021	AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCTGAC TCCGCTGTTCCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTCGGACTG
7081	TGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCCCTGGATGCCAGGGAC ACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTG
7141	TCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCCTGTTTCTGGA AGGGGCGTCCAACGGGGCGGACGGGGTCGCGATGACCGTTTACGCCGGGGACAAAGACCT
	7167 ECO47III
7201	GCTGCTTGGGAACCACGCGCAGTGCCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCT CGACGAACCCTTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGGCGA
7261	GCGAGCTGCGGTCACCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGT CGCTCGACGCCAGTGGGGTCGTCGGCCACAGACACGGGCCCTCTTCGGGGTCCCGAGACA
7321	GGCGGCCCCGAGGAGGAGACACAGACCCCCGTCGCCTGGTGCAGCTGCTCCGCCAGCA CCGCCGGGGGCTCCTCCTGTGTCTGGGGGCAGCGGACCACGTCGACGAGGCGGTCGT
7381	CAGCAGCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCT
7441	AGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCAT TCCGGAGACCCCGAGGTCCGTGTTGCTTGCGGCGAAGGAGTCCTTGTGGTTCTTCAAGTA
7501	CTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCG GAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCCTCGACTGCACCTTCTACTCGCACGC
	**********
7561	GGACTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCCAGG CCTGACGCGAACCGACGCGTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGGTCC
	7575 FSP1
	************************************
7621	CCCCAGAGCTGAATGCAGTAGGGGGCTCAGAAAAGGGGGCAGGCA
	* * * * * * * * * * * * * * * * * * * *
7681	GTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACGGTGCAGGGTAGCAGTGCACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGCCAC
7741	**> ATCGAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATA TAGCTCCAGCTGAGATCTCCTAGGGGCCCATGGCTCGAGCTTAAGCATTAGTACCAGTAT
	7747 SAL1

FIG. 21 (CONTINUED)

	30/34
6121	**************************************
·	Topo_II_cleavage_s
	:::::::::::::::::::::::::::::::::::::::
	NFkB ========
	<pre>Intron1 ************************************</pre>
6181	GGAACCAGCGACATGCGGAGAGCAGCGCAGCGACTCAGGGCGCTTCCCCCGCAGGTGTCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCGCTGAGTCCCGCGAAGGGGGCGTCCACAG
	ite
6241	CTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGCGC
6301	CGTGCTGGCCTTCGGCTTCGCGCTGGACGGGCCCCGGGGGGCCCCCCGAGGCCTTGCACGACCGGAAGCCGAAGCGCGACGACCTGCCCCGGGCGCCCCCGGGGGGGG
6361	CACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGGGTGGTGGTCGCACGCGTCGATGGACGGGTTGTGCCACTGGCTGACGCCCCCTCGCC
	6372 FSP1
6421	GGCGTGGGGGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGCACGCCCCCGACGACGACGACGCGCGCG
6481	CTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCGCGCGCG
6541	GTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCCGAAGCATGGTCGAGCCGCGACGGTGAGTCCGGGCCGGGGGGGG
6601	GCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCTGGGCCGAGACCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCC
6661	CCTGCCAGCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCGGACGGTCGGGGCCCACGCTCCTCCGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGG
6721	CAAGAGGCCCAGGCGTGGCGCTGCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTCGTTCTCCGGGTCCGCACCGCGACGGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAG
6781	CTGGGCCCACCCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTGTG
6841	TGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTC

FIG. 21 (CONTINUED)

6901 CCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCCATCCACATCGCGGCCACCACG

5461	CCCTCGGGTTACCCCACAGCCTAGGC GGGAGCCCAATGGGGTGTCGGATCCG			
		*	Sp1	
5521	GGCGTCCCTGCACCCTGGGAGCGCGACCGCGGGACGTGGGACCCTCGCGCT			
5581	CCGGGTCCGCCCGGAAGCAGCTGCGC GGCCCAGGCGGGCCTTCGTCGACGCG			
			o_II_cleavage_si ******	te
5641	GGGCACAGACGCCCAGGACCGCGCTT CCCGTGTCTGCGGGTCCTGGCGCGAA			
			2F ****	
5701	CGTCCTGCCCCTTCACCTTCCAGCTC GCAGGACGGGGAAGTGGAAGGTCGAG			
	<u>.</u>	•	,	E ****
5761	CCCTTCCCAGGTCCCGGCCCAGCCCCGGGAAGGGTCCAGGGCCGGGTCGGGG	TTCCGGGCCC BAAGGCCCGGG	TCCCAGCCCCTCCCCT AGGGTCGGGGAGGGGA	TCCTTTTC AGGAAAAG
	Sp1 =======			
	2F ****	NFkB *****	*****	h *****
5821	CGCGGCCCGCCCTCTCCTTCGCGGCGCGCGGGGGGGGGG			
			5860	5875
			ECO47III	FSP1
	TRT5'		ECO47111	FSP1
5881	TRT5'  ***********************************	CCCGCGATGC	:CGCGCGCTCCCCGCTG	CCGAGCCG
5881 5941	**************************************	CCCGCGATGC GGGCGCTACG	CGCGCGCTCCCCGCTG GCGCGCGAGGGGCGAC	CCGAGCCG GGCTCGGC
•	**************************************	CCCGCGATGC GGGCGCTACG	CGCGCGCTCCCCGCTG GCGCGCGAGGGGCGAC	CCGAGCCG GGCTCGGC
5941	CGTGGGAAGCCCTGGCCCCGGCCACCGCACCCTTCGGGACCGGGGCCGGTGCCTGCC	CCCGCGATGC GGGGCGCTACG CCGCGAGGTGC GGCGCTCCACG	CCGCGCGCTCCCCGCTG GCGCGCGCGAGGGCGAC CTGCCGCTGGCCACGTT GACGGCGACCGGTGCAA	CCGAGCCG CGTGCGGC CGTGCGCCG
5941 6001	CGTGGGAAGCCCTGGCCCCGGCCACCGCCACCGCCACCCTTCGGGACCGGGCCGGTGCCACCACTACACGCGAGGGACGACGACGACGAGGACGACGACGAGGACGAC	CCCGCGATGC GGGGCGCTACG CCGCGAGGTGC GGCGCTCCACG GGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	CCGCGCGCTCCCCGCTG GCGCGCGCGAGGGCGAC CTGCCGCTGGCCACGTT CACGCGACCGGTGCAA GGGGACCCGGCGGCTTT CCCTGGGCCGCCCCAAA	CCGAGCCG CGTGCGGC CGTGCGCCG CCACGCCGC CCGCGCGCGC CGCGCGCGCGCG

FIG. 21 (CONTINUED)

	FIG. 21
5401	GCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCCAGGGCCTCCACATCATGGCCCCTCGGTCGCCGGTTTCCCAGCGGCGTGCGT
5341	GTGCCCGGAGCCCGACGCCCCGCGTCCGGACCTGGAGCCAGGCCCTGGGTCTCCGGATCAGCCCCGGGCCTCGGGCCTCCGGGCCCTAGTCCACGGGCCTCCGGGACCCAGAGGCCTAGTC
5281	TTCGTCCCCAWGCCGCGTCTACGCGCCTYCCGTCCTCCCCTTCACGTTCCGGCATTCGTGAAGCAGGGGTWCGGCGCAGATGCGCGGARGGCAGGAGGGGAAGTGCAAGGCCGTAAGCAC
5221	CAGGGAGGCACTCCGGGGAGGTCCSGCGTGCCCGTCCAAGGGAGCAATGCGTCCTTCGGGGTCCCTCCGTGAGGCCCCTCCAGGSCGCACGGGCAGGTTCCCTCGTTACGCAGGAAGCCC
5161	AACGTGCAAAGAGGAAATGAAGGGCCTGKGTCAAGGAGCCCAAGTNGGCGGGGRAGTTTG TTGCACGTTTCTCCTTTACTTCCCGGACMCAGTTCCTCGGGTTCANCCGCCCCYTCAAAC
5101	GGGGATRAACAGANTNGGGGGKGGTWGGGTTNAKGGTGGGAACNCCTTNGCGCCTGGAG CCCCTAYTTGTCTNANCCCCCMCCAWCCCAANTMCCACCCTTGNGGAANCGSCGGACCTC
5041	NGCCANGRAGGGGCCAGGTTCCAANTTCCCAACCKTTTTWGGARGGACNGCCCCCAGGG NCGGTNCYTCCCCCGGTCCAAGGTTNAAGGGTTGGMAAAAWCCTYCCTGNCGGGGGTCCC
4981	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
4921	ИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИ
4861	ИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИ
4801	ИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИ
4741	ИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИ
4681	ИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИ
4621	ИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИИ
4561	ИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМ
4501	ИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМ
4441	ИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМ
4381	ИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМИМ
4321	GGTGTTTTTAAGCCAATNANAAAATTTTTTNATGTTGTTTNNNNNNNNNN

$\cdot$	
1 ATCGATTGGGCCCGAGATCTCGCGCGCGGGGCCTGCCATGGGACCCACTGCAGGGGC TAGCTAACCCGGGCTCTAGAGCGCGCGCTCCGGACGGTACCCTGGGTGACGTCCCCG	
3615 3636 BGL2 NCO1	
1 TGGGANGCTGCAGGCTTCAGGTCCCAGTGGGGTTGCCATCTGCCAGTAGAAACCTGA ACCCTNCGACGTCCGAAGTCCAGGGTCACCCCAACGGTAGACGGTCATCTTTGGACT	
AGAATCAGGGCGCGAGTGTGGACACTGTCCTGAATCTCAATGTCTCAGTGTGTGCTG TCTTAGTCCCGCGCTCACACCTGTGACAGGACTTAGAGTTACAGAGTCACACACGAC	
1 CATGTAGAAATTAAAGTCCATCCCTCCTACTCTACTGGGATTGAGCCCCTTCCCTATGTACATCTTTAATTTCAGGTAGGGAGGATGAGATGACCCTAACTCGGGGAAGGGATA	
1 CCCCCAGGGGCAGAGGAGTTCCTCTCACTCCTGTGGAGGAAGGA	
	****
1 TTTCACTGCTGGTACTGAATCCACTGTTTCATTTGTTGGTTTGTTT	
***********	
1 AGCGGTTTCACTCTTGTTGCTCAGGCTGGANGGAGTGCAATGGCGCGATCTTGGCTT TCGCCAAAGTGAGAACAACGAGTCCGACCTNCCTCACGTTACCGCGCTAGAACCGAA	ract
ALU	
************	
1 GCAGCCTCTGCCTCCCAGGTTCAAGTGATTCTCCTGCTTCCGCCTCCCATTTGGCTG CGTCGGAGACGGAGGGTCCAAGTTCACTAAGAGGACGAAGGCGGAGGGTAAACCGAC	GGA CCT
*******	****
1 TTACAGGCACCCGCCACCATGCCCAGCTAATTTTTTGTATTTTTAGTANANACNGGG AATGTCCGTGGGCGGTGGTACGGGTCGATTAAAAAACATAAAAATCATNTNTGNCCC	GTG CCAC
	. A
======================================	CANC
LU	
======================================	
1 TGCCTCTGCCTCTAAAATTGCTGGGATTACAGGTGTNANCCACCATGCCCAACTCAACGGAGACGGAGGATTTTAACGACCCTAATGTCCACANTNGGTGGTACGGGTTGAGT	
1 TTTACTCTGTTTANAAACATCTGGGTCTAAGGTAGGAANCTCACCCCACTCAATTTTAAAATGAGACAAATNTTTGTAGACCCAGATTCCATCCTTNGAGTGGGGTGAGTTAAAA	

760 his gly his val arg lys ala phe lys ser his val leu arg pro CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA val pro gly asp pro ala gly leu his pro leu his ala ala leu GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG 790 gln pro val leu arg arg his gly glu gln ala val cys gly asp CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT 800 807 ser ala gly arg ala ala pro ala phe gly gly OP TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTTGGT GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA GTATGCTGCGTGAACTTGCGGAAGACAGTGGTGAACTTCCCTGTAGAAGACGAGGC CCTGGGTGGCACGCTTTTGTTCAGATGCCGGCCCACGGCCTATTCCCCTGGTGCGGCCT GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTC CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA ACTCTTTGGGGTCTTGCGGCTGAAGTGTCACAGCCTGTTTCTGGATTTGCAGGTGAACAG CCTCCAGACGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA CGCATGTGTGCTGCAGCTCCCATTTCATCAGCAAGTTTGGAAGAACCCCACATTTTTCCT GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG GATGTCGCTGGGGGCCAAGGGCGCCGCCGGCCTTCTGCCCTCCGAGGCCGTGCAGTGGCT GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACGTGCCACTCCT GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG CTACGTCCCAGGGAGGGGGGGCGCCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG GCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACCTGCCGTCTTCACTTCCC CACAGGCTGGCGCTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTC CACTCCCCACATAGGAATAGTCCATCCCCAGATTCGCCATTGTTCACCCCTCGCCCTGCC CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT GGGGGTCCCTGTGGGTCAAATTGGGGGGGGGGTGCTGTGGGAGTAAAATACTGAATATATG 

FIG. 20 (CONTINUED)

550 ala lys phe leu his trp leu met ser val tyr val val glu leu GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG 560 leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC 580 arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ile gly ile arg gln his leu lys arg val gln leu arg glu leu ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG 610 ser glu ala glu val arg gln his arg glu ala arg pro ala leu . TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG 640 pro ile val asn met asp tyr val val gly ala arg thr phe arg CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC arg glu lys arg ala glu arg leu thr ser arg val lys ala leu AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG 670 phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu TTC AGC GTG CTC AAC TAC GAG CGG GCG CGC CCC GGC CTC CTG gly ala ser val leu gly leu asp asp ile his arg ala trp arg GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC 700 thr phe val leu arg val arg ala gln asp pro pro pro glu leu ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG 710 tyr phe val lys val asp val thr gly ala tyr asp thr ile pro TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC 730 gln asp arg leu thr glu val ile ala ser ile ile lys pro gln CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG 740 asn thr tyr cys val arg arg tyr ala val val gln lys ala ala AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

FIG. 20 (CONTINUED)

340 leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA 350 leu ser ser leu arg pro ser leu thr gly ala arg arg leu val CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG 370 glu thr ile phe leu gly ser arg pro trp met pro gly thr pro GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC 380 arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC 400 leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG 410 val leu leu lys thr his cys pro leu arg ala ala val thr pro GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA 430 ala ala gly val cys ala arg glu lys pro gln gly ser val ala GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG 440 ala pro glu glu glu asp thr asp pro arg arg leu val gln leu GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG 460 leu arg gln his ser ser pro trp gln val tyr gly phe val arg CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG 470 480 ala cys leu arg arg leu val pro pro gly leu trp gly ser arg GCC TGC CTG CGC CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG 490 his asn glu arg arg phe leu arg asn thr lys lys phe ile ser CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC 500 510 leu gly lys his ala lys leu ser leu gln glu leu thr trp lys CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG 520 met ser val arg asp cys ala trp leu arg arg ser pro gly val ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT 530 gly cys val pro ala ala glu his arg leu arg glu glu ile leu GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

FIG. 20 (CONTINUED)

	•							120						
tyr TAC	leu CTG	pro CCC	asn AAC	thr ACG	val GTG	thr ACC	asp GAC	130 ala GCA	leu CTG	arg CGG	gly GGG	ser AGC	gly GGG	ala GCG
trp TGG	gly GGG	leu CTG	140 leu CTG	leu CTG	arg CGC	arg CGC	val GTG	gly GGC	asp GAC	asp GAC	val GTG	leu CTG	150 val GTT	his CAC
leu CTG	leu CTG	ala GCA	arg CGC	cys TGC	ala GCG	leu CTC	phe TTT	160 val GTG	leu CTG	val GTG	ala GCT	pro CCC	ser AGC	cys TGC
ala GCC	tyr TAC	gln CAG	170 val GTG	cys TGC	gly GGG	pro CCG	pro CCG	leu CTG	tyr TAC	gln CAG	leu CTC	gly GGC	180 ala GCT	ala GCC
thr ACT	gln CAG	ala GCC	arg CGG	pro CCC	pro CCG	pro CCA	his CAC	190 ala GCT	ser AGT	gly GGA	pro CCC	arg CGA	arg AGG	arg CGT
leu CTG	gly GGA	cys TGC	200 glu GAA	arg CGG	ala GCC	trp TGG	asn AAC	his CAT	ser AGC	val GTC	arg AGG	glu GAG	210 ala GCC	GGG gly
val GTC	pro CCC	leu CTG	gly GGC	leu CTG	pro CCA	ala GCC	pro CCG	220 gly GGT	ala GCG	arg AGG	arg AGG	arg CGC	gly GGG	gly
ser AGT	ala GCC	ser AGC	230 arg CGA	ser AGT	leu CTG	pro CCG	leu TTG	pro CCC	lys AAG	arg AGG	pro CCC	arg AGG	240 arg CGT	gly GGC
ala GCT	ala GCC	pro CCT	glu GAG	pro CCG	glu GAG	arg CGG	thr ACG	250 pro CCC	val GTT	gly GGG	gln CAG	gly GGG	ser TCC	trp TGG
ala GCC	his CAC	pro CCG	gly GGC	arg AGG	thr ACG	arg CGT	gly GGA	pro CCG	ser AGT	asp GAC	arg CGT	gly GGT	270 phe TTC	cys TGT
	val GTG													glu GAG
gly GGT	ala GCG	leu CTC	290 ser TCT	gly GGC	thr ACG	arg CGC	his CAC	ser TCC	his CAC	pro CCA	ser TCC	val GTG	300 gly GGC	arg CGC
gln CAG	his CAC	his CAC	ala GCG	gly GGC	pro CCC	pro CCA	ser TCC	310 thr ACA	ser TCG	arg CGG	pro CCA	pro CCA	arg CGT	pro CCC
trp TGG	asp GAC	thr ACG	320 pro CCT	cys TGT	pro CCC	pro CCG	val GTG	tyr TAC	ala GCC	glu GAG	thr ACC	lys AAG	330 his CAC	phe TTC

FIG. 20 (CONTINUED)

MetSerValTyrValValGluLeuLeuArgSerPhePhe TyrValThrGluThrThrPheGlnLysAsnArgLeuPhe PheTyrArgLysSerValTrpSerLysLeuGlnSerIle GlyIleArgGlnHisLeuLysArgValGlnLeuArgGlu LeuSerGluAlaGluValArgGlnHisArgGluAlaArg ProAlaLeuLeuThrSerArgLeuArgPheIleProLys ProAspGlyLeuArgProIleValAsnMetAspTyrVal ValGlyAlaArgThrPheArgArgGluLysArgAlaGlu ArgLeuThrSerArgValLysAlaLeuPheSerValLeu AsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAla SerValLeuGlyLeuAspAspIleHisArgAlaTrpArg ThrPheValLeuArgValArgAlaGlnAspProProPro GluLeuTyrPheValLysValAspValThrGlyAlaTyr AspThrIleProGlnAspArgLeuThrGluValIleAla SerIleIleLysProGlnAsnThrTyrCysValArgArg TyrAlaValValGlnLysAlaAlaHisGlyHisValArg LysAlaPheLysSerHisValLeuArgProValProGly AspProAlaGlyLeuHisProLeuHisAlaAlaLeuGln ProValLeuArgArgHisGlyGluGlnAlaValCysGly AspSerAlaGlyArgAlaAlaProAlaPheGlyGly

GCAGCGCTGCGTCCTGCCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCG	1 met CG ATG
pro arg ala pro arg cys arg ala val arg ser leu leu ar CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG	rg ser GC AGC
20 his tyr arg glu val leu pro leu ala thr phe val arg ar CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CC	30 rg leu GC CTG
gly pro gln gly trp arg leu val gln arg gly asp pro al GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GG	la ala CG GCT
phe arg ala leu val ala gln cys leu val cys val pro tr TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TG	
ala arg pro pro pro ala ala pro ser phe arg gln val se GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TC	er cys CC TGC
80  leu lys glu leu val ala arg val leu gln arg leu cys g.  CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GA	90 lu arg AG CGC
gly ala lys asn val leu ala phe gly phe ala leu leu as GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG GA	sp gly AC GGG
ala arg gly gly pro pro glu ala phe thr thr ser val ar GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CC	

MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPOGWRLVORGDP AAFRALVAOCLVCVPWDARPPPAAPSFRQVSCLKELVARVLQRL CERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPNTVTDALR-GSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG ARRRGGSASRSLPLPKRPRRGAAPEPERTPVGOGSWAHPGRTRG PSDRGFCVVSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPP STSRPPRPWDTPCPPVYAETKHFLYSSGDKEQLRPSFLLSSLRP SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLEL LGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPQGSVAAPEE EDTDPRRLVQLLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNE RRFLRNTKKFISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC VPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTETTFQKNR LFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPAL LTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKA LFSVLNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPP ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ KAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI EQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPQGSI LSTLLCSLCYGDMENKLFAGIRRDGLLLRLVDDFLLVTPHLTHA KTFLRTLVRGVPEYGCVVNLRKTVVNFPVEDEALGGTAFVQMPA HGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR NMRRKLFGVLRLKCHSLFLDLOVNSLOTVCTNIYKILLLQAYRF HACVLOLPFHOOVWKNPTFFLRVISDTASLCYSILKAKNAGMSL GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGSLRTAQ TQLSRKLPGTTLTALEAAANPALPSDFKTILD

### FIG. 17

TTATGTCACGGAGACCACGTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTG GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGG AGCCAGAACGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACT GTTCAGCGTGCTCAACTACGAGCGGGCGCGCGCCCCGGCCTCCTGGGCGCCTCTGTGCT GGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCCAGGA CCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC CCAGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAAACCCCAGAACACGTACTGCGT GCGTCGGTATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAG CCACGTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCT GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGGATTCGGCGGGACGGGC TGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACCTCACCCACGCGAAAA CCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACTTGCGGA AGACAGTGGTGAACTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGA TGCCGGCCCACGGCTATTCCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGG TGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC GCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGT GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTC ATCAGCAAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCATCTCTGACACGGCCTCCC TCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGCCAAGGGCCCCG CCGGCC7TCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCAAGCATTCCTGCTCAAGC TGACTCGACACCGTGTCACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGC AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGCAGCCAACCCGG CACTGCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCG GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGC TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC AGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC CCCAGATTCGCCATTGTTCACCCCTCGCCCTGCCCTCTTTGCCTTCCACCCCCACCATC CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTGGGG AAAAAAAAAAAAA

1 gcagcgctgc gtcctgctgc gcacgtggga agccctggcc ccggccaccc ccgcgatgcc 61 gcgcgctccc cgctgccgag ccgtgcgctc cctgctgcgc agccactacc gcgaggtgct 121 gccgctggcc acgttcgtgc ggcgcctggg gccccagggc tggcggctgg tgcagcgcg 181 ggacccggcg gctttccgcg cgctggtggc ccagtgcctg gtgtgcgtgc cctgggacgc 241 acggccgccc cccgccgccc cctccttccg ccaggtgtcc tgcctgaagg agctggtggc 301 ccgagtgctg cagaggctgt gcgagcgcgg cgcgaagaac gtgctggcct tcaacttcac 361 gctgctggac ggggcccgcg ggggcccccc cgaggccttc accaccagcg tgcgcagcta 421 cctgcccaac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc tgctgctgcg 481 ccgcgtgggc gacgacgtgc tggttcacct gctggcacgc tgcgcgctct ttgtgctggt 541 ggctcccagc tgcgcctacc aggtgtgcgg gccgccgctg taccagctcg gcgctgccac 601 tcaggcccgg cccccgccac acgctagtgg accccgaagg cgtctgggat gcgaacgggc 661 ctggaaccat agcgtcaggg aggccggggt ccccctgggc ctgccagccc cgggtgcgag gaggcgcggg ggcagtgcca gccgaagtct gccgttgccc aagaggccca ggcgtggcgc 781 tgcccctgag ccggagcgga cgcccgttgg gcaggggtcc tgggcccacc cgggcaggac 841 gcgtggaccg agtgaccgtg gtttctgtgt ggtgtcacct gccagacccg ccgaagaagc 901 cacctetttg gagggtgege tetetggeae gegeeaetee eacceateeg tgggeegeea 961 gcaccacgeg ggcccccat ccacategeg gccaccacgt ccctgggaca cgccttgtcc 1021 cccggtgtac gccgagacca agcacttcct ctactcctca ggcgacaagg agcagctgcg 1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga ggctcgtgga 1141 gaccatettt etgggtteca ggeeetggat geeagggaet eccegeaggt tgeeecgeet 1201 gccccagcgc tactggcaaa tgcggcccct gtttctggag ctgcttggga accacgcgca 1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgctg cgagctgcgg tcaccccagc 1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggcccccg aggaggagga 1381 cacagaccc cgtcgcctgg tgcagctgct ccgccagcac agcagcccct ggcaggtgta 1441 cggcttcgtg cgggcctgcc tgcgccggct ggtgccccca ggcctctggg gctccaggca 1501 caacgaacgc cgcttcctca ggaacaccaa gaagttcatc tccctgggga agcatgccaa 1561 gctctcgctg caggagctga cgtggaagat gagcgtgcgg gactgcgctt ggctgcgcag 1621 gagcccaggg gttggctgtg ttccggccgc agagcaccgt ctgcgtgagg agatcctggc 1681 caagttcctg cactggctga tgagtgtgta cgtcgtcgag ctgctcaggt ctttctttta 1741 tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga gtgtctggag 1801 caagttgcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc gggagctgtc 1861 ggaagcagag gtcaggcagc atcgggaagc caggcccgcc ctgctgacgt ccagactccg 1921 cttcatcccc aagcetgacg ggctgcggcc gattgtgaac atggactacg tcgtgggagc 1981 cagaacgttc cgcagagaaa agagggccga gcgtctcacc tcgagggtga aggcactgtt 2041 cagegtgete aactaegage gggegeggeg eeeeggeete etgggegeet etgtgetggg 2101 cctggacgat atccacaggg cctggcgcac cttcgtgctg cgtgtgcggg cccaggaccc 2161 gccgcctgag ctgtactttg tcaaggtgga tgtgacggc gcgtacgaca ccatcccca 2221 ggacaggete aeggaggtea tegecageat cateaaacee cagaacaegt aetgegtgeg 2281 toggtatgoo gtggtocaga aggoogooca tgggcacgto cgcaaggoot toaagagooa 2341 cgtctctacc ttgacagacc tccagccgta catgcgacag ttcgtggctc acctgcagga 2401 gaccageceg etgagggatg cegtegteat egageagage teeteeetga atgaggeeag 2461 cagtggcctc ttcgacgtct tcctacgctt catgtgccac cacgccgtgc gcatcagggg 2521 caagteetae gteeagtgee aggggateee geagggetee atecteteea egetgetetg 2581 cagcetgtge tacggegaca tggagaacaa getgtttgeg gggattegge gggaeggget 2641 gctcctgcgt ttggtggatg atttcttgtt ggtgacacct cacctcaccc acgcgaaaac 2701 cttcctcagg accctggtcc gaggtgtccc tgagtatggc tgcgtggtga acttgcggaa 2761 gacagtggtg aacttccctg tagaagacga ggccctgggt ggcacggctt ttgttcagat 2821 gccggcccac ggcctattcc cctggtgcgg cctgctgctg gatacccgga ccctggaggt 2881 gcagagegae tactecaget atgeceggae etceateaga gecagtetea cetteaaceg 2941 cggcttcaag gctgggagga acatgcgtcg caaactcttt ggggtcttgc ggctgaagtg 3001 tcacagecty tttetggatt tgeaggtgaa cageetecag aeggtgtgea ceaacateta 3061 caagateete etgetgeagg egtacaggtt teaegeatgt gtgetgeage teeeatttea 3121 tcagcaagtt tggaagaacc ccacattttt cctgcgcgtc atctctgaca cggcctccct 3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg ctgggggcca agggcgccgc 3241 cggccctctg ccctccgagg ccgtgcagtg gctgtgccac caagcattcc tgctcaagct 3301 gactcgacac cgtgtcacct acgtgccact cctggggtca ctcaggacag cccagacgca 3361 gctgagtcgg aagctcccgg ggacgacgct gactgccctg gaggccgcag ccaacccggc 3421 actgecetea gaetteaaga eeateetgga etgatggeea eeegeeeaca geeaggeega 3481 gagcagacac cagcagecet gtcacgcegg getetacgte ecagggaggg aggggeggee 3541 cacacccagg cccgcaccgc tgggagtctg aggcctgagt gagtgtttgg ccgaggcctg 3601 catgteegge tgaaggetga gigieegget gaggeetgag egagtgteea geeaaggget 3661 gagtgtccag cacacctgcc gtcttcactt ccccacaggc tggcgctcgg ctccacccca 3721 gggccagett tteeteacea ggageeegge tteeacteee cacataggaa tagteeatee 3781 ccagattege cattetteae ecctegeeet geeeteettt geetteeaee eecaeeatee 3841 aggtggagac cctgagaagg accctgggag ctctgggaat ttggagtgac caaaggtgtg 3901 ccctgtacac aggcgaggac cctgcacctg gatgggggtc cctgtgggtc aaattggggg 3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa aaaaa

### FIG. 15 (CONTINUED)

3960 818	4020 838	4089 848	4149 868	4209 888	4274 903	4339 917	4401 935	4468 946	4528 966	4588 986	4665 989
TGT	999	A TCG	TCT S	AAA K	ACG G gtgagtacttattttaactaga T D	TTG AAT GTT ATT GGA AGA AATT TGG AAA AAG TTG GCC L N V I G R K I W K K L A	AAA TG gtacgtgtc K W	AAA K	ACT T	ATA I	$\mathtt{TAA}$ tgtcattttcaatttattatatacatcctttattactggtgtcttaaacaatattattactaagtata $\star$
GCA A	ATG M	ag ,	AAT N	TTA L	taa	TT L	ytac	TTG	${ m TTG}$	AGA R	taa
TTA L	CAT H	aato	$\mathrm{TTC}_{\mathrm{F}}$	TAC Y	attt	AAC K	T. S	GGT	TCA S	AGA R	ıttacı
${ m TTG}$	AAA K	aaat	AAA K	GCA A	actt	X X	AAA K	GAT D	CAG Q	CAT H	atte
ACA T	ACG I	tgac	TCA S	CAA GCA TAC 1	gagt	TGG W	GTC V	AGA R		TTA L	ıcaat
GAT D	CTG L	tago	AAT N	GCA	G gt D	ATT	GAA E	ATG M	CAA Q	$ ext{TTT}$	taaa
$_{\rm L}^{\rm CTT}$	GĄG E	ataa	CAC H	AGA R	ACG T	X X	GCA (	GGA G	TAC	TTA L	gtct
TCT S	GTA V	ctga	ACC T	ATG 1	ATA I	AGA R	TCT	$_{\rm L}^{\rm CTT}$	ATA I	GTG V	tggt
AGG R	TCT S	gtaa	ATT I	TGT /	TTC	GGA G	TCC	TGT C	CTA L	CAG Q	ıttac
ATG M	ACA T	ctgt	GAC	ATG M	ATG M	ATT	TTG L	$ ext{TTT}$	CAG Q	CGA R	ttta
AAC	TCT S	gtatactgtgtaactgaataatagctgacaaataatcag A	ATT I	TCT S	AGA R	GTT V	TTC F	$_{\rm L}^{\rm CTT}$	GAA E	${ m TTG}$	atco
GTG V	AAC N	AG R	TTT F	TAC Y	CAA	AAT N	CGT R	g G	TTC F	GTT V	atac
TCT S	TTT F	CTA	GTA V	GGA	Oμ	TTG	AGG R	atca	TGC	CCA P	ıttat
$_{ m F}^{ m TTC}$	TTA L	ATT I	CAA	CTA L	ATT I	CTT	AGT S	acac	CCA P	AGA R	ıttta
GGT	GCC	AAA K	GCA	AGG R	TTT F	g AT	ACG T	attg	:AT I	CTA L	tcaa
$\mathrm{TTC}_{\mathrm{F}}$	GAA	TAC	TTT F	TAT Y	ATA I	ctta	TAT Y	caat	TAT Y	CCG P	attt
$ ext{TTC}$	GAT D	TTT F	TCC	ATA I	GAT D	taac	GGA TAT ACG P	tcag	AAA TAT C	AAG K	tgtc
CCA P	ATT I	$ ext{TTT}$	GCA A	AAT N	AAG K	taat	ATA TTA ( I L	ggtctcgagacttcagcaatattgacacatcag G	TTC F	ATC	TAA *
ATG M	AAA K	TCT S	CTT L	TGC C	ATG	tcat	ATA I	tcga	TCT	$_{\rm L}^{\rm CTT}$	GAT
AGA R	CCT	AAA K	AGC	TGC	AGG R	aaagtcattaattaaccttag AT CTT ' L	GAA	ggtc	CCC	GAT D	GCT
3901 799	3961 819	4021 839	4090	4150 869	4210	4275 904	4340 918	4402 936	4469 947	529 967	4589
3	n	4	4	4	4	4	4	4	4	Þ	4

FIG. 15 (CONTINUED)

3155 591	3215 611	3275 631	3343 643	3405 659	3465 679	3532 692	3593 708	3653 728	3713 748	3777 764	3840 778	3900 798
89 tatataatgcgcgattcctcattattaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 82	56 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 92 K S C Y D R I K Q D L M F R I V K K K L	216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3 612 K D P E F V I R K Y A T I H A T S D R A 6	276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttatttttcattggaattttttaacaa 3 632 T K N F V S E A F S Y F	344 attettttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3 644 D M V P F E K V V Q L L S M K T 6	406 TCA GAT ACT TTG GTT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3 660 S D T L F V D F V D Y W T K S S S E I F 6	466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgttgaattgtaataaca 3 680 K M L K E H L S G H I V K	A GTT GGT ATC CCT CAG GGC TCA V G I P Q G S	594 ATT CTG TCA TCT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3 709 I L S S F L C H F Y M E D L I D E Y L S 7	654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3729 F T K K G S V L L R V V D D F F L F I T 7	3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgtcattcc $3749$ V N K K D A K K F L N L S L R G	778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3 765	841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3779 I N F F N E S K K 7
30	31	33	m T	m ~	è ¯	m ¯	m T	m ·	m ·	m	m	m ¯

F/G. 15 (CONTINUED)

2336 375	2396 395	2465 405	2525 425	2585 445	2645 465	2705 485	2775 495	2835 515	2906 524	2967 542	302 <i>7</i> 562	3088 581
A ACT T	AAG K	GGA	GCG A	TAT Y	AAA K	GAG	A ACT T	ACG T	gtattaatttttggtcatcaatgtactttacttctaatctatta	GTG V	GAG E	gtaat
GAA	ATA I	CTT	TTT F	TTT F	TGG ₩	AAC N	s AAA K	ATT I	atc	CCT P	$_{\rm L}^{\rm TTG}$	ဗ္ဗ ဗ
CTC	AAC N	GTC V	ATA I	TTT F	ATT I	ATA I	CAG	CTC	tctā	CGA R	AAC N	TTT F
g AC	AGT S	CTA L	CAA Q	TCT S	GAT D	AAA K	ACT T	CGT R	tact	TTA L	$ ext{TTT}$	ATG M
gtattgtataaaatttattaccactaacgattttaccag	ATG M	TGG W	AAG K	CAA Q	AAA	GAA E	GAT D	$ ext{TTT}$	actt	ACT T	CCA P	CGA R
ttt	TTA L	GAA E	CGC R	TTA L	AGA A	TTT F	A ATG	ACC T	aatgt	CAA Q	ATT I	CAC H
ıacga	$_{\rm Y}^{\rm TAT}$	ATT I	AAA K	ATT I	TTT F	GCG A	r AGG R	AAT N	satca	AAC N	GGT G	AAG K
acta	CAT H	GAA E	GAG E	CCT P	$_{\rm Y}^{\rm TAT}$	GAA E	r GTT V	AAG K	ggto	ACG	AGT S	$_{\rm L}^{\rm CTT}$
taco	TTA L	TCA S	TTT F	ATA I	GTT V	ATG M	N N	AAG K	ttt	AGT S	AGT S	CTT L
ttat	AGT S	ATT I	GAT D	ATA I	ACT T	AAA K	J AAC N	CCT P	taat	GTC V	GAA E	GAT D
laaat	$ ext{TTT}$	cag	AGT S	$ ext{TTT}$	CGA R	ATG M	tcaç	TTA L		TTA L	GAA E	AAG K
ıtata	TCT	ıcaat	TTA L	TCG S	AAT N	TCA S	attt	CTA L	AAG K	ATG M	AAT N	AAG K
attg	GAG E	ıttaa	TGC	AAT N	CGA R	ACA T	taat	CGT R	ATA I	AAA K	ATC I	TTT F
G gt D	TAC Y	ccaaattttttaccattaattaacaatcag	ATG M	TAC Y	TTA L	ATT I	aagtatttttgcaaaaagctaatatttcag	ATT I	TTA L	AAA K	TTA L	ACT T
AAA K	AGA R	acca	AAA K	$_{\rm L}^{\rm CTA}$	GAT D	$ ext{TTT}$	ycaaa	GTT V	TTC F	AAC N	CAT H	CTT L
$_{\rm L}^{\rm TTA}$	TCG S	ttt	GCG A	TGG W	AGT S	CCC	ttt	GCA A	AGA R	TCA S	AAA K	CTT L
ATA I	TTA	ıattt	AAT N	TAC Y	TCA S	CGA R	attt	CCA P	AAA K	GGT G	CTG L	AAG K
ATA I	AAA K	Jccae	TCA S	ATC I	GAA E	TGC C	aaagt	CCT	AGA R	ATG M	ATA I	ATG M
GAG E	TTG L	tatg	AGG R	$\mathrm{TTC}$	ACT	TTG L	ttta	TTG L	TTA L	agcag	TCG S	TAC Y
$ ext{TTT}$	${ m TTC}$	gtaa	AAA K	GAA E	ATC I	$_{ m L}^{ m CTC}$	gtatt	ACT T	AAT	ttag	GCA	GTT V
2268 366	2337	2397 396	2466 406	2526 426	2586 446	2646 466	2706 486	2776 496	2836 516	2907 525	2968 543	3028 563

FIG. 15 (CONTINUED)

1529 148	1601 155	1661 175	1721 195	1781 215	1841 235	1907 245	1967 265	2027 285	2087 305	2147 325	2207 345	2267 365
CAA Q	299 3	GAC D	GTG V	AAA K	TAT Y	AAC N	AGG R	GTA V	ACA T	ATT I	GCG A	ATA I
TGG W	ATC I	AAT N	ACT	CGC R	TCC	TTT A	CCA P	$_{\rm L}^{\rm CTG}$	CAA	TAT Y	TTT F	AGG R
AAT N	gtaaataccggttaagatgttgcgcactttgaacaagactgacaagtatag T ATC I	CCA P	GAA E	GCC A	TCA S	TAT TT Y	TTT F	CCA P	GAA E	CCA P	GTG V	CAA Q
AAA K	agtat	$_{\rm L}^{\rm CTT}$	GAG E	AGC S	AGG R	CTA T	ATT I	ATT	ATT I	TGC	CAG Q	AAC N
TCA S	yacaé	GCT A	$ ext{TTT}$	AAA K	TAC	e	TGG W	GTG V	TTA L	TAT Y	AAC	GGT G
GAG E	yactç	GAG E	GTG V	AAT N	$ ext{TTT}$	tag i	CAA	AAA K	CCT P	CAT H	CCG P	TGG W
CTT L	асрас	$ ext{TTT}$	AAT N	CAA Q	ATT I	G gtaactaatactgttatccttcataactaattttag D	$_{\rm L}^{\rm CTT}$	CAC H	TAC Y	AAC N	AAG K	ATC I
ATA I	tgaë	ATT I	AAT N	ACT T	AGC S	acta	TGG W	TTG L	GTA V	TAC Y	TTA L	TTA L
TCT S	actt	AGT S	AAA K	ATT I	$ ext{TTT}$	cata	ATG M	CAA Q	AAG K	GTT V	TCC	AAA K
ATA I	gege	GGA G	TTT F	TCC	AGG R	sctto	CAC H	AAG K	CTA L	AAA K	$_{\rm Y}^{\rm TAT}$	CCT P
CTT L	atgtt	AAA K	$_{\rm L}^{\rm CTT}$	ACA T	AGT S	tato	GTA V	GTG V	CTC	TCA S	AGT S	TTT F
TAC	taage	TCC	CCA P	GAA E	ATT I	actgi	ACA T	CĄA Q	CGT R	CTA L	$_{\rm L}^{\rm CTT}$	GTG V
AAT	ggtt	TTA L	ATA I	ATT I	TCA S	taati	AAC N	$_{\rm F}^{\rm TT}$	AAA K	TCT S	ATC I	CGA R
CCT P	atace	TTA L	0 0	ACC T	ATT I	tàac	CGG R	GCA A	CCC P	ATT I	AAA K	GTT V
TTT F	gtaaa	TAC	TCT S	CGA R	AGC S			AAC N	GTG V	CGT R	GAA E	CTT
ACT T	AT	CAT H	ATT I	AAG K	AAT N	CAA Ó	TGT C	ATA I	GTT V	CAT H ·	GAT D	ATT I
TCT S	GAA	ATG M	CAG Q	AGA R	TGG W	AAG K	ATT I	CTT L	ACA T	$\overset{\text{CTC}}{\text{L}}$	GAT D	TCC
GTT V	TTA L	GCC	$_{\rm L}^{\rm CTT}$	AAA K	TCC	TTT F	TCT	GGA G	AGT S	CGA R	CAC H	CGA R
$^{\mathrm{CTC}}_{\mathrm{L}}$	TTG	GAT D	TAC Y	AAA K	GTT V	AAG K	CAC H	$\frac{\text{TTT}}{\text{F}}$	CAG Q	AAG K	ACC T	$_{\rm L}^{\rm CTT}$
GAT D	CTT	AGT S	AAT N	TCA S	GAA E	AAG K	TTA L	caa o	TCA S	GCA A	GAC D	$ ext{TT}$
1470 129	1530 149	1602 156	1662 176	1722 196	1782 216	1842 236	1908 246	1968 266	2028 286	2088 306	2148 326	2208 346

FIG. 15 (CONTINUED)

	œ	œ	œ	æ	7	7	2	6
80 1160 2240 3240 4400 7240 9880 958	1018 20	1078 40	1138 60	1198	127 86	1332	140 113	1469 128
c d t c c c c c c c c c c c c c c c c c	GTA V	TCG	CTT L	CCA P	cag	CTA L	aga	AAT N
accaag troott attro agata atatt gaaaa agcat taatg tttaatg tttaatg	₽ S				999	F.	aag	A A A
	, TAT Y	AGC S	TTT F	TCT	tat	AAT N	tca	CAA
tattattet	CAA	GCA A	ATT I	TCT S	aata	AGG 7	gtt	GTA V
aaattttccccccccccccccccccccccccccccccc	AAT N	CCG P	TCT S	TTT F	gcta	CGA	tact	GGA G
ct c	GAG E	TCG S	TTT F	CAA	yata	CGT.	caat	AAC
tcctcgaacgctcctaaatctctggaaatattttta gttattagtgatcgataatatttctattttatcggt aagacttttactttattaatttacttttcaaatata acttctagccaaccgcgtgtttctaccccgtcattg tgatgagactattagattcattacagtccgtcat tgatgagactattagattcattacagtccgtcat aatccgcgaaagttttttgattcattacagtccgtct taacgcggaaagttttttctatttctattccatg taacgcggaatcgtacctttttcatttctattcccc aaagattggtgattctactctttttcactattccccc aaagattggtgattctactctttttactattagt aaaaaatcctaaattataaaatattagt catatatagttaataaaatattaagtcaatattg	CTA (	000 0	TCC	GTT (	cgg	TTT GAG CGT F E R	ctg	GTA AAC (
a ta a a a a a a a a a a a a a a a a a	TTT C F	AGA G R	ACG T	GGT G	tatt	TT G	ttac	CAT G
	Ö F	Ğ R			tto	F.	tat.	ATG C
ncgo trac lcta lcta lgaa lgaa lgta ltat	cGC R	rTG L	CAA Q	GAA	tt	A AG	Jaaa	TA :
	CTT L	GTT V	GTA V	GAT D	gatt	GAA	tgtç	GCC A
CC transcription of the control of t	ATT I	$_{\rm L}^{\rm CTT}$	GAT D	CCA P	tt	GAT D	aat	CGA R
goct cota gocta gocta cotota tott tott	AGG R	CAA Q	AGC S	AAG K	ttgt	$\mathrm{TTC}_{\mathbf{F}}$	ttct	$ ext{TTT}$
aatt cccc aaccc catc tccc actc actc accc	AGC S	GTA V	AGA R	AGT S	attt	ATG TTC GAT GAA AGT M F D E S	gtaaggtattctaattgtgaaatatttacctgcaattactgtttcaaagag	GAT TTT CGA GCC D F R A
a de trada a a trada a a de trada a a de trada a a a trada a a tra	AAA	TAT	TTG	GAC	atat	CAG	gtaa	GAA
intractticctticticataagctaattgcticctcgaacgctcctaaatctictggaaatatttitacaaga acaatacaaga acaataccaattaggacgtta aggaggaccaaataccaagtccaattggaggtgtiattagggaggattattacttictattictattitatcggtcgttagaggaccaaacgaccaacggcggtgtticctaatticctatticaaattatticggtatticattic	CCC 7	GAT 7 D.	CGC 7	TTC (F	gtatatattttttttttttttttttttctattcgggatagctaatatgggcag	AAA K	ATG O	CAT (H
aaaataaataataataataataataataataataataat	ACC C	AAT G N	GAA C	GGC T	GAG 9	GTA P	TCC P	AAT C
ttt agga cagt cagt ttc tta tta	CAT A H							
t caa aaaa aaaaa caaaaaa caaaaaaaaaaaaaa		TTA L	TGC C	GTC V	TCA S	GTT V	TTT F	aaa
a c t t t t t t t t t t t t t t t t t t	CAC H	ACC T	ATA I	GTA V	CAG Q	AAT N	999 9	cgat
tttta aaagg gaagt tttat atte aatte aatte caaatte	GAA E	TGT C	AAT N	ACT T	TCA S	GCG A	AAA K	aaccgataaag
ccgaacaaacaaaaaaaaaaaaaaaaaaaaaacctaaccccacccccc	ACC T	CTA L	AGC S	TCG S	TGC C	ATA I	ATG M	attt
aggtaccaaacaaacaaacaaacaaacaaacaaacaaaca	ATG M	TAC	TAT	CAT	AAA K	CTA	CTG	ttgtatt
8 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	59 A	21 )		39 61	99 81	73	33	06 14
10044506C88	σ	10	1079	11	11	12	13	14

2401	CCCTGAAAAT	CCAAATGTTA	ATCTTCTAAT	GAGACTTACA	GATGACTATC
2451	TTTTGATTAC	AACTCAAGAG	AATAATGCAG	TATTGTTTAT	TGAGAAACTT
2501	ATAAACGTAA	GTCGTGAAAA	TGGATTTAAA	TTCAATATGA	AGAAACTACA
2551	GACTAGTTTT	CCATTAAGTC	CAAGCAAATT	TGCAAAATAC	GGAATGGATA
2601	GTGTTGAGGA	GCAAAATATT	GTTCAAGATT	ACTGCGATTG	GATTGGCATC
2651	TCAATTGATA	TGAAAACTCT	TGCTTTAATG	CCAAATATTA	ACTTGAGAAT
2701	AGAAGGAATT	CTGTGTACAC	TCAATCTAAA	CATGCAAACA	AAGAAAGCAT
2751	CAATGTGGCT	CAAGAAGAAA	CTAAAGTCGT	TTTTAATGAA	TAACATTACC
2801	CATTATTTTA	GAAAGACGAT	TACAACCGAA	GACTTTGCGA	ATAAAACTCT
2851	CAACAAGTTA	TTTATATCAG	GCGGTTACAA	ATACATGCAA	TGAGCCAAAG
2901	AATACAAGGA	CCACTTTAAG	AAGAACTTAG	CTATGAGCAG	TATGATCGAC
2951	TTAGAGGTAT	CTAAAATTAT	ATACTCTGTA	ACCAGAGCAT	TCTTTAAATA
3001	CCTTGTGTGC	AATATTAAGG	ATACAATTTT	TGGAGAGGAG	CATTATCCAG
3051	ACTTTTTCCT	TAGCACACTG	AAGCACTTTA	TTGAAATATT	CAGCACAAAA
3101	AAGTACATTT	TCAACAGAGT	TTGCATGATC	CTCAAGGCAA	AAGAAGCAAA
3151	GCTAAAAAGT	GACCAATGTC	AATCTCTAAT	TCAATATGAT	GCATAGTCGA
3201	CTATTCTAAC	TTATTTTGGA	AAGTTAATTT	TCAATTTTTG	TCTTATATAC
3251	ጥርርርርርጥጥጥጥር	CCCTTTTCCC	GTTTTGGGG		

### FIG. 13 (CONTINUED)

1	MEVDVDNQAD	NHGIHSALKT	CEEIKEAKTL	YSWIQKVIRC	RNQSQSHYKD
51	LEDIKIFAOT	NIVATPRDYN	EEDFKVIARK	EVFSTGLMIE	LIDKCLVELL
101	SSSDVSDRQK	LQCFGFQLKG	NQLAKTHLLT	ALSTQKQYFF	QDEWNQVRAM
151	IGNELFRHLY	TKYLIFQRTS	EGTLVQFCGN	NVFDHLKVND	KFDKKQKGGA
201	ADMNEPRCCS	TCKYNVKNEK	DHFLNNINVP	NWNNMKSRTR	IFYCTHFNRN
251	NQFFKKHEFV	SNKNNISAMD	RAQTIFTNIF	RFNRIRKKLK	DKVIEKIAYM
301	LEKVKDFNFN	YYLTKSCPLP	ENWRERKQKI	ENLINKTREE	KSKYYEELFS
351	YTTDNKCVTQ	FINEFFYNIL	PKDFLTGRNR	KNFQKKVKKY	VELNKHELIH
401	KNLLLEKINT	REISWMQVET	SAKHFYYFDH	ENIYVLWKLL	RWIFEDLVVS
451	LIRCFFYVTE	QQKSYSKTYY	YRKNIWDVIM	KMSIADLKKE	TLAEVQEKEV
501	EEWKKSLGFA	PGKLRLIPKK	TTFRPIMTFN	KKIVNSDRKT	TKLTTNTKLL
551	NSHLMLKTLK	NRMFKDPFGF	AVFNYDDVMK	KYEEFVCKWK	QVGQPKLFFA
601	TMDIEKCYDS	VNREKLSTFL	KTTKLLSSDF	WIMTAQILKR	KNNIVIDSKN
651	FRKKEMKDYF	RQKFQKIALE	GGQYPTLFSV	LENEQNDLNA	KKTLIVEAKQ
701	RNYFKKDNLL	QPVINICQYN	YINFNGKFYK	QTKGIPQGLC	VSSILSSFYY
751	ATLEESSLGF	LRDESMNPEN	PNVNLLMRLT	DDYLLITTQE	NNAVLFIEKL
801	INVSRENGFK	FNMKKLQTSF	PLSPSKFAKY	GMDSVEEQNI	VQDYCDWIGI
851	SIDMKTLALM	PNINLRIEGI	LCTLNLNMQT	KKASMWLKKK	LKSFLMNNIT
901	HYFRKTITTE	DFANKTLNKL	FISGGYKYMQ	CAKEYKDHFK	KNLAMSSMID
951	LEVSKIIYSV	TRAFFKYLVC	NIKDTIFGEE	HYPDFFLSTL	KHFIEIFSTK
1001	KYIFNRVCMI	LKAKEAKLKS	DQCQSLIQYD	A	

FIG. 14

### 12/34 .

	•				
1	AAAACCCCAA	AACCCCAAAA	CCCCTTTTAG	AGCCCTGCAG	TTGGAAATAT
51	AACCTCAGTA	TTAATAAGCT	CAGATTTTAA	ATATTAATTA	CAAAACCTAA
101	ATGGAGGTTG	ATGTTGATAA	TCAAGCTGAT	AATCATGGCA	TTCACTCAGC
151	TCTTAAGACT	TGTGAAGAAA	TTAAAGAAGC	TAAAACGTTG	TACTCTTGGA
201	TCCAGAAAGT	TATTAGATGA		CTCAAAGTCA	TTATAAAGAT
251	TTAGAAGATA	TTAAAATATT	TGCGCAGACA	AATATTGTTG	CTACTCCACG
301	AGACTATAAT	GAAGAAGATT	TTAAAGTTAT	TGCAAGAAAA	GAAGTATTTT
351	CAACTGGACT		CTTATTGACA		TGAACTTCTT
401	TCATCAAGCG	ATGTTTCAGA	TAGACAAAAA	CTTCAATGAT	TTGGATTTCA
451	ACTTAAGGGA	AATCAATTAG	CAAAGACCCA	TTTATTAACA	GCTCTTTCAA
501	CTCAAAAGCA	GTATTTCTTT	CAAGACGAAT	GGAACCAAGT	TAGAGCAATG
551	ATTGGAAATG	AGCTCTTCCG	ACATCTCTAC	ACTAAATATT	TAATATTCCA
601	GCGAACTTCT	GAAGGAACTC	TTGTTCAATT	TTGCGGGAAT	AACGTTTTTG
651	ATCATTTGAA	AGTCAACGAT	AAGTTTGACA	AAAAGCAAAA	AGGTGGAGCA
701	GCAGACATGA	ATGAACCTCG	ATGTTGATCA	ACCTGCAAAT	ACAATGTCAA
751		GATCACTTTC		CAACGTGCCG	AATTGGAATA
801	ATATGAAATC	AAGAACCAGA	ATATTTTATT	GCACTCATTT	TAATAGAAAT
851	AACCAATTCT	TCAAAAAGCA	TGAGTTTGTG	AGTAACAAAA	ACAATATTTC
901	AGCGATGGAC		CGATATTCAC	GAATATATTC	AGATTTAATA
951	GAATTAGAAA		GATAAGGTTA	TCGAAAAAAT	TGCCTACATG
1001	CTTGAGAAAG	TCAAAGATTT	TAACTTCAAC		CAAAATCTTG
1051	TCCTCTTCCA	GAAAATTGGC	GGGAACGGAA	ACAAAAAATC	GAAAACTTGA
1101	TAAATAAAAC	TAGAGAAGAA		ACTATGAAGA	GCTGTTTAGC
1151	TACACAACTG		CGTCACACAA	TTTATTAATG	AATTTTTCTA
1201	CAATATACTC	CCCAAAGACT	TTTTGACTGG	AAGAAACCGT	AAGAATTTTC
1251	AAAAGAAAGT	TAAGAAATAT	GTGGAACTAA	ACAAGCATGA	ACTCATTCAC
1301	AAAAACTTAT		GATCAATACA		CATGGATGCA
1351	GGTTGAGACC		ATTTTTTATTA		GAAAACATCT
1401	ACGTCTTATG	GAAATTGCTC	CGATGGATAT	TCGAGGATCT	CGTCGTCTCG
1451	CTGATTAGAT	GATTTTTCTA	TGTCACCGAG	CAACAGAAAA	GTTACTCCAA
1501	AACCTATTAC	TACAGAAAGA	ATATTTGGGA	CGTCATTATG	AAAATGTCAA
1551	TCGCAGACTT	AAAGAAGGAA	ACGCTTGCTG	AGGTCCAAGA	AAAAGAGGTT
1601	GAAGAATGGA	AAAAGTCGCT		CCTGGAAAAC	TCAGACTAAT
1651	ACCGAAGAAA		GTCCAATTAT	GACTTTCAAT	AAGAAGATTG
1701	TAAATTCAGA	CCGGAAGACT	ACAAAATTAA		GAAGTTATTG
1751	AACTCTCACT		GACATTGAAG	AATAGAATGT	TTAAAGATCC
1801	TTTTGGATTC		ACTATGATGA	TGTAATGAAA	AAGTATGAGG
1851	AGTTTGTTTG	CAAATGGAAG		AACCAAAACT	CTTCTTTGCA
1901	ACTATGGATA		ATATGATAGT	GTAAACAGAG	AAAAACTATC
1951	AACATTCCTA	AAAACTACTA		TTCAGATTTC	TGGATTATGA
2001	CTGCACAAAT		AAGAATAACA	TAGTTATCGA	TTCGAAAAAC
2051	TTTAGAAAGA	AAGAAATGAA		AGACAGAAAT	TCCAGAAGAT
2101		GGAGGACAAT		ATTCAGTGTT	CTTGAAAATG
2151	AACAAAATGA		AAGAAAACAT	TAATTGTTGA	AGCAAAGCAA
2201	AGAAATTATT		TAACTTACTT	CAACCAGTCA	TTAATATTTG
2251	CCAATATAAT	TACATTAACT		GTTTTATAAA	CAAACAAAAG
2301	GAATTCCTCA		GTTTCATCAA	TTTTGTCATC	ATTTTATTAT
2351		AGGAAAGCTC			
2001	,				

FIG. 13

> NFkB\_CS1 GGGRQTYYQC NFkB-MHC-I.2 TGGGCTTCCCC

### Intron1

301 GCTGGGGTTGAGGGCGGCGGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGCGACTC CGACCCCAACTCCCGCCGGCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

NFkB\_CS1 GGGRQTYYQC NFkB\_CS2 RGGGRMTYYCC

Topo\_II\_cleavage\_site RNYNNCNNGYNGKTNYNY \*\*\*\*\*\*

361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGGTCCCCGCGAAGGGGGCGTCCACAGGACGACTTCCTCGACCACCGGGCTCACGACGTCTCC

## Telomerase Specific Motifs

MOTIF T'	E V	EAEVR	ENNVR	EKEVE	ENINC
_		13	12	12	0
	Н	Igi	FI	ISI	ÞΕΙ
	W J	<b>WSKLQS1</b>	[WKLLCR]	IWDVIMK	LWNKLITE
	y Rk w	FFYRKS\	'VYFRKD	LYYYRKN.	.VYFRHD1
MOTIF T	Wl FFY TE	546 WLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGI 13 EAEVR	429 WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFI 12 ENNVR	441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSI 12 EKEVE	Sc_Est2 366 WLFRQLIPKIIQTFFYCTEISSTVT.IVYFRHDTWNKLITPFI 9 ENNVC
		546	429	441	366
	TRT con	hTRT	SPTRT	Ea_p123	Sc_Est2

## Telomerase RT Motifs (Fingers)

MOTIF B'	Y d GipQGs 1S 1 Y	PELYFVKVDVTGAYDTI 104 YVQCQGIPQGSILSTLLCSLCY	99 YLQKVGIPQGSILSSFLCHFYM	67 PKLFFATMDIEKCYDSV 117 YKQTKGIPQGLCVSSILSSFYY	85 YIREDGLFQGSSLSAPIVDLVY	hPQG pP hh h	
MOTIF A	p lyF D cYD i	69 PELYFVKVDVTGAYDTI	66 RKKYFVRIDIKSCYDRI	67 PKLFFATMDIEKCYDSV :	68 PELYFMKFDVKSCYDSI	h hDh AF h	۸۵
MOTIF 2	fr I	0 LRPIV	0 FRLIT	O FRPIM	2 FRIIA	hR h	
MOTIF 1	R iPKk	11 SRLRFIPKPDG 0 LRPIV		10 GKLRLIPKKTT 0 FRPIM	13 SKMRIIPKKSN 2 FRIIA	p hh h K	
	TRT con	htrr	SPTRT	Ea_p123	Sc Est2	RT con	

# Telomerase RT Motifs (Palm, Primer Grip)

		192	176	174	141		
MOTIF E	wgs l	WCGLLLDTRTL	FFGFSVNMRSL	VSRENGFKFNMKKLQT 28 WIGISIDMKTL	WKHSSTMNNFH	hLG h	
		24	22	28	25		
		ŢW	TVI	LQT	ILA	ㅁ	
Ω	n X	ILRK	LEK	MKK	RDK	Gh h cK	
MOTIF D	되	NY.	STS	KFN	KAN	h 1	
MOJ		EYGC	KHNE	ENGE	KYN2	ਠ	
	p	GVP	GFE	VSR	GFQ		
		15	15	15	15		
MOTIF C	111rl DDfL it	15 LLLRLVDDFLLVT 15 GVPEYGCVVNLRKTVV 24 WCGLLLDTRTL	16 VLLRVVDDFLFIT 15 GFEKHNFSTSLEKTVI	24 LLMRLTDDYLLIT	18 LILKLADDFLIIS 15 GFQKYNAKANRDKILA 25	h Y DDhhh	ţz
		15	16	24	18		
	TRT con	hTRT	SPTRT	Ea_p123	Sc_Est2	RT_con	

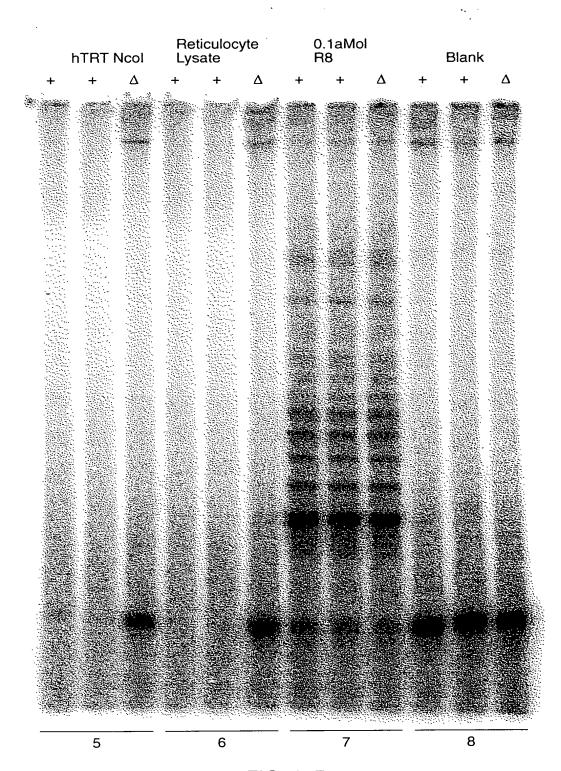


FIG. 10B

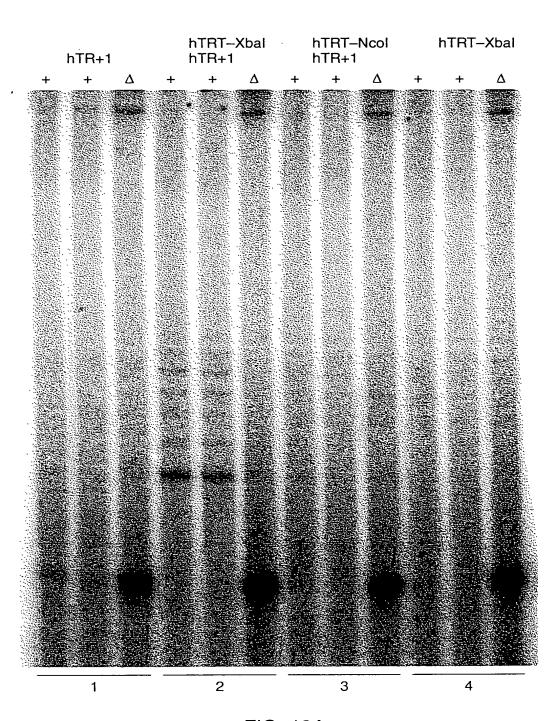
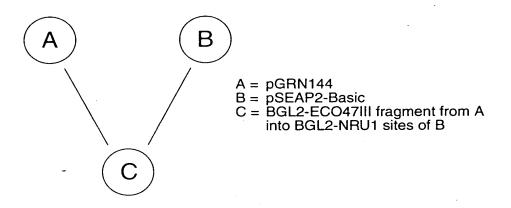


FIG. 10A



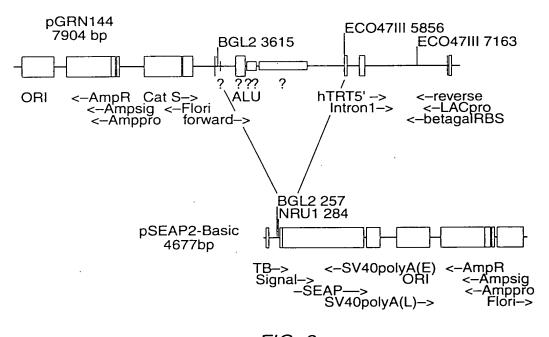


FIG. 9

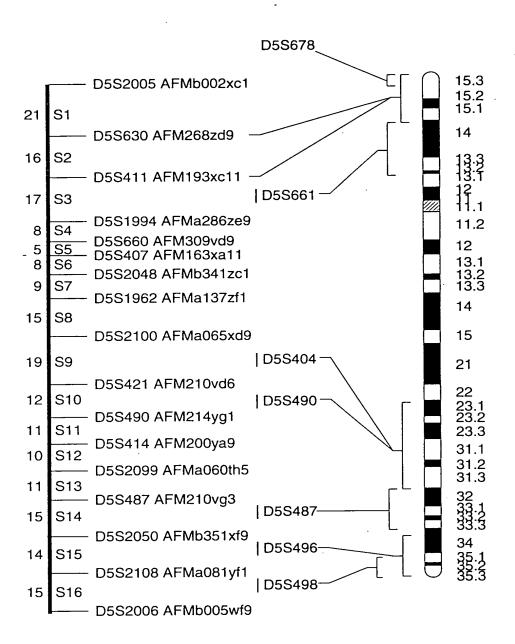


FIG. 8

TOWNSEND & TOWNSEND & CREW 15389-26PC 1215 WOODWORKS 408 378-2790

### 5/34

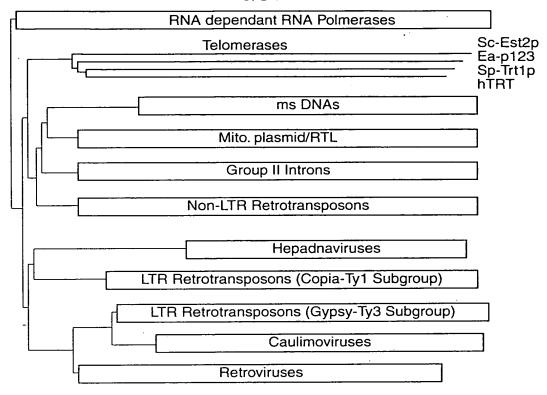


FIG. 6

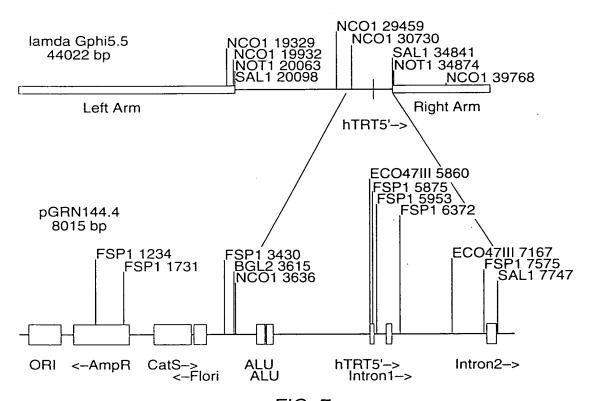


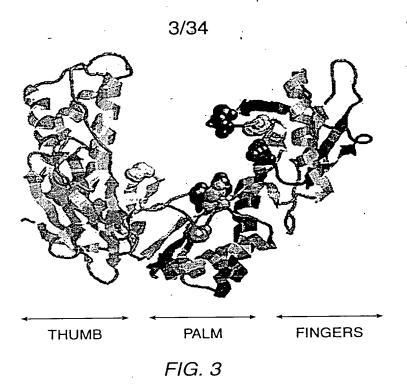
FIG. 7

WCLIFT P Y RK W L h I K 429 WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFITSMKM 8 546 WLMSVYVVELLRSFTYTFQKNRLFFYRKSVWSKLQSIGIRQHLK 10 441 WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIWDVIMKMSIADLKK 8 366 WLFRQLIPKIIQTFFYCTEISSTVT-IVYFRHDTWNKLITPFIVEYFK 8

TRT con Sp\_rrtlp hrrr Ea\_p123 Sc\_Est2p

•	•
4/34	181 197 179 146
82 87 100 68 68 25	
Motif 1  h hrhipkk p frhi h h k  cveewkkslgfapgklrlipkkrut o frlitaipcrgadeeffivkenhknalop crnhnsytlsnenkrkningerkengenitkengengengengengengengengengengengengenge	KY Q GIPQGS LS hL h Y DL F LLRL DDFLhIT A F h G c p N CK W G S  SQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFT 6 LLRVUDDFLFITVNKKD 0 AKFENLSLRGFEKHNFSTSLERTVI 17 KKRMPFFGFSV KSYVQCQGIPQGSILSTLCSLCYGDMENKLFAGI 5 LLRLVDDFLLTTQENN 0 AVLFIEKLINVSRENGFKFNMKRLQT 23 QDYCDWIGISI KFYKQTKGIPQGLCVSSILSSFYYATLEESSLGFL 14 LMRLTDDYLLITTQENN 0 AVLFIEKLINVSRENGFKFNMKRLQT 23 QDYCDWIGISI KFYKQTKGIPQGLCVSSILSFYSEFK 8 ILKLADDFLISTDQQQ 0 VINIKKLAMGGFQKYNAKANRDKTLA 20 KELEVWKHSST F GAPARLGLPQGSLISPILCNIVMTLVDNWLEDYI 55 YVRYADDILIGVLGSKN 2 KMIKRDLNNFLNS-LGLTMNEEKTLI 4 ETPARFLGYNI GIRYQYNVLPQGWKGSPAIFQSSMTKILEPFKKQN 4 IYQYMDDLYVGSDLEIG 1 HRTKIEELRQHLLRWGLTTPDKKHQK 0 EPPFLWMGITL
TRT con Sp_rrtlp hrrr Ea_p123 Sc_Est2p RT con Sc_a1 Dm_rarr HIV-1	TRT con Sp_Trt1p hTRT Ea_p123 Sc_Est2p RT con Sc_a1 Dm_TART

F1G. 4



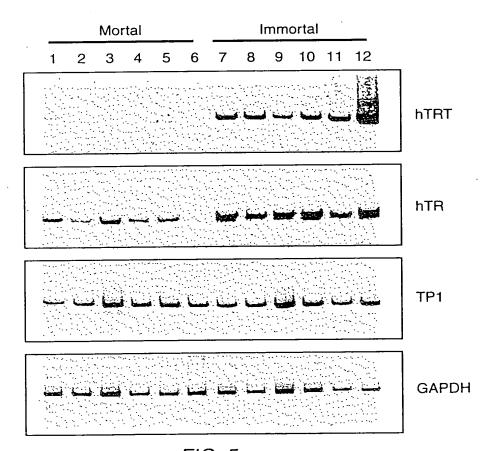


FIG. 5

pl 10.6 11.3 10.1		
Mol. weight 116.000 127.000 123.000 103.000		
Sp_Trt1p - hTRT - Ea_p123 Sc_Est2p		
T 1.2 A B'CD E	msDNAs Mito.plasmid/RTL	Hepadnaviruses  LTR Retrotransposons (Copia-Ty1) <b>EXEM-ZZACED-LEGIZ-D</b> LTR Retrotransposons (Gypsy-Ty3) <b>EXEM-ZZICETANMIZED</b> Caulimoviruses  Retroviruses
A 50 aa		

MOLII U AKFLHWLMSVYVVELLRSFFYVTETTFQKNR ISEIEWLVLGKRSNAKMCLSDFEKRKQIFAEFIYWLYNSFIIPILQSFFYITESSDLRNR LKDFRWLFISDIWFTKHNFENLNQLAICFISWLFRQLIPKIIQTFFYCTEISSTVT- TREISWMQVET-SAKHFYYFDHEN-IYVLWKLLRWIFEDLVVSLIRCFFYVTEQQKSYSK	MOLIÍ 1 LFFYRKSVWSKLQSIGIRQHLKRVQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGL TVYFRKDIWKLLCRPFI-TSMKMEAFEKINENNVRMDTQK-TTLPPAVIRLLPKKNTF IVYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLSNFNHSKMRIIPKKSNNEF TYYYRKNIWDVIMKMSI-ADLKKETLAEVQEKEVEEWKKS-LGFAPGKLRLIPKKTTF	MOLIÍ 2 RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLNYERA RLITN-LRKRFLIKMGSNKKMLVSTNQTLRPVASILKHLINEESSGIPFNLEVYMKLLTF RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKEF RPIMTFNKKIVNSDRKTTKLTNTKLLNSHLMLKTLKN-RMFKDPFGFAVFNYDDVMKKY * *	MOLIÍ 3 (A) KKDLLKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPEFVIRKYATIHATS KQRLLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN EEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKN
human tez1 EST2 p123	human tez1 EST2 p123	human tez1 EST2 p123	tez1 EST2 p123